AD-A210 107

(2)

SSI Project # 1146

OPTIMAL SENSOR CONFIGURATION AND SURVIVABLE PROCESSING WITH CORRELATED NOISE

Office of the Naval Research Center

Department of the Navy

Contract No. N00014-88-C-0736

Final Report

May 1, 1989

Shahjahan Mahmood

Raman K. Mehra

Scientific Systems, Inc.
One Alewife Place
Cambridge, MA 02140

TABLE OF CONTENTS

LIST OF FIGURES

FORWARD

This report describes the work performed by Scientific Systems, Inc. from 10/15/88 to 4/30/88 under the Navy Award No. N00014-88-C-0736. The work represents the first of a three-phase project aimed at development of theory, computational techniques and a commercial grade software for solving decentralized estimation problems in the presence of correlated sensor noise. Solutions are available for the correlated sensor noise only In this report, the available results have been extended in many directions and the distributed tracking problem has been solved with correlated sensor measurement noise The objective of Phase I effort was to demonstrate the feasibility of a solution technique of this problem. The problem scenario will be broadened and the results will be analyzed in detail in Phase II.

Dr. Keith Bromley was the project director for the Navy. Dr. Shah Mahmood was the project manager and principal investigator. Dr. Raman K. Mehra was the project supervisor. Excellent typing and reliable documentation by Mimi Starr is greatly appreciated.

SUMMARY

The research proposed under this project is concerned with the problem of accurate target tracking in the SDI environment. The SDI scenario considered in this report is as follows. Multiple sensors that are possibly located at dispersed geographical locations are observing a single target. The measurement noises are correlated across the sensors. The local processors attached with the sensors possibly may not have the complete information of the target model The tracking problem can be solved in two ways: first, the sensor measurements can be transmitted to a central node where the optimal state can be estimated conditional on all the measurement history - this is centralized processing Alternatively, each sensor can process its own measurement locally and send (some function of) processed data to the central node Here it is fused with other incoming data into the global estimate - this is decentralized processing. The latter is superior to the centralized processing from the system performance considerations in the face of node failures

There exists a good amount of literature in the area of decentralized estimation in presence of uncorrelated sensor noise. In this work, the available results have been extended to the case of correlated sensor noise under the assumption that the local processors have the knowledge of the target model. Each local agent needs to send the central node only one vector of the dimension of the state. The available results for the uncorrelated sensor noise have been extended in other directions too. For example, it has been established that there is a tradeoff between the computational requirements at a node and the transmission requirements from this node Using this analysis and by assigning the major computational load to the local nodes, the complexity of the central node has been reduced to an adder.

The decentralized estimation problem has been analyzed both in discrete time and in continuous time domain. It has been proposed that the nonlinear measurement equations

can be dealt with by a recent technique known as "Modified Gain Extended Kalman Filter (MGEKF) more efficiently than the conventional Extended Kalman Filter (EKF)

CHAPTER 1

INTRODUCTION AND SUMMARY

1.1 Introduction and Overall Objective of Phase I

This report contains results from the investigations of the project "Optimal Sensor Configuration and Survivable Processing with Correlated Noise." This is a Phase I effort sponsored by the Naval Research Center of Arlington, Virginia through Small Business Innovative Research (SBIR) Program under contract #N00014-88-C-0736. This job was sponsored by the Navy through the Strategic Defense Initiative Organization (SDIO). This work was performed at Scientific Systems, Inc. (SSI), Cambridge, Massachusetts during the period October, 1988 - April, 1989.

The surveillance, acquisition and tracking function for the Strategic Defense Initiative (SDI) includes sensing information for battle management and processing signals and data for discrimination of threatening reentry vehicles from other objects. As each potential reentry vehicle is released from its post-boost vehicle, it begins ballistic midcourse flight accompanied by deployment hardware and possibly by decoys. Each credible object must be accounted for in a birth-to-death track, even if the price is many decoy false alarms. Interceptor vehicles of the defense must also be tracked.

The research proposed under this project is concerned with the problem of accurate target tracking in the SDI environment. The problem of tracking boosters and reentry vehicles has received much attention in the past. However, the problem has generally been structured as a single-sensor, single-target problem. The SDI environment is much more complex, however, in that many targets must be tracked simultaneously. In addition, many sensors may be used to track a single target. Although the use of multiple sensors

intuitively allows more accurate and robust tracking to be attained, it is not yet clear how to fully exploit this additional capability. In addition, the SDI environment is such that the sensor observations are subject to noises which are correlated across sensors This further complicates the problem.

Ideally, the problem of target tracking in a multi-sensor environment with limited communication between the sensors is a distributed estimation and detection problem which uses no centralized signal processing. The information available to each sensor consists of its own measurement history and possibly all past messages received from other sensors In this report we shall refer to this problem also as a distributed tracking problem. The overall objective of this 3-phase project is to obtain a practical solution of the distributed tracking algorithm for multiple targets under the assumption that the measurement noises are correlated across the sensors. This report contains Phase I activities and indicates the feasibility of the proposed technique for solving this problem. Detailed development of the technique will be carried out in the forthcoming Phase II and finally, these ideas will be commercialized in Phase III.

The overall problem of distributed tracking is indeed a very difficult and complex one This problem can be decomposed into various subproblems that can be arranged into hierarchical layers according to their complexity as shown in Figure 1.1. Obviously, the starting point is the centralized processing where raw measurements from all the sensors are transmitted to a common fusion center for simultaneous processing - this is the lowest layer of the hierarchy.

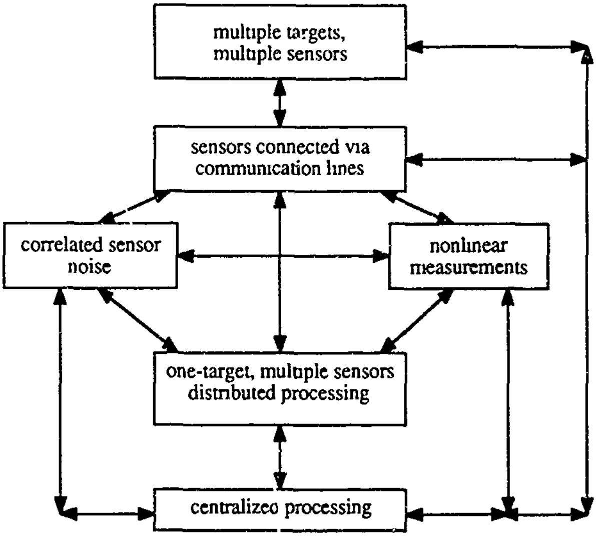


Figure 1.1: Decomposition of the Target Tracking Problem in an SDI Framework

One of the serious criticisms of this technique is that the system is vulnerable to catastrophic failure if the central node fails. On the other hand, in the case of distributed processing, the observed data from each sensor is processed locally. The processed data are then communicated to a centralized coordinator who forms the global estimate. In this case, the system performance will have a graceful degradation in the event of sensor/node failure. This is the primary motivation for opting for the distributed processing over the centralized one.

Various components of the distributed tracking problem are shown in Figure 1.1. This figure indicates a systematic approach to the overall solution of the problem. Any subset of the blocks of this figure which are connected by arrows can be combined to form a meaningful tracking problem. The core of the distributed tracking problem is the case of

one-target, multiple sensor with distributed processing" and is shown at the second layer of Figure 1 1 In this scenario, time delays due to computation and communication processes are neglected It is assumed that the sensors are communicating with each other instantaneously in a broadcast mode Although this is a hypothetical setting, the analysis of this case exhibits the structure of the decentralized processing problem. In a realistic situation the measurements from various sensors will be correlated primarily because the sensors are observing the same target through the same atmospheric medium. The decentralized estimation problem then becomes considerably difficult if the correlation is taken into account. The measurement equations which are nonlinear in states also adds further complication to this problem. Once the distributed tracking problem is solved for correlated noise and nonlinear measurement equations, the effect of a realistic communication channel can then be considered In an SDI framework, direct communication takes place only between adjacent satellites arranged in a ring structure about the Earth During this phase of analysis, time delays due to computation and communication processes must be taken into account if the algorithm is intended for practical application Finally, the scenario for multiple targets can be analyzed as an extension of the one-target case.

Because of the time limitations in Phase I, we could not solve all subproblems of Figure 1 1 We have explored only the basic problem. the case of correlated measurement noise and nonlinear measurement equations as shown in layers 2 and 3 of our decomposition

1.2 A Simplified Generic SDI Tracking Problem

Let us consider an example of a simple SDI tracking problem with one target only - this scenario will be broadened to multiple targets in Phase II. Formulation of the decentralized estimation problem will be motivated through this example. Consider then a

generic SDI tracking problem with one target in a ballistic trajectory and three satellite sensors as shown in Figure 1.2. The sensors are assumed to make time-synchronized measurements. The measurements are taken through one or more of the laser, infrared and radar sensors that are located at appropriate places of the orbiting satellite stations.

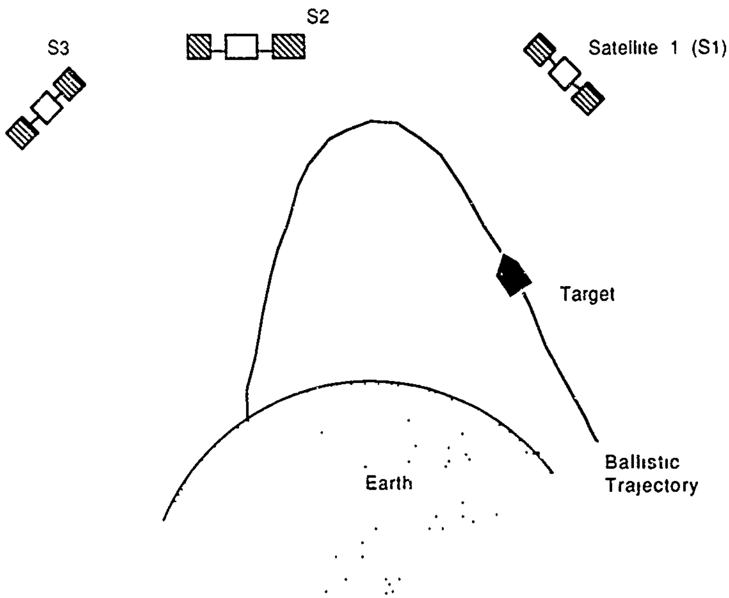


Figure 1.2: A Simplified SDI Tracking Problem

Both active and passive sensors will be considered under the scope of the present task. An active sensor such as a radar measures the range (R), azimuth angle (A) and elevation angle (E) of the target in a reference frame that is centered at the radar. On the other hand, a passive sensor such as an infrared sensor produces "bearings only"

measurements The SDI tracking problem can be formulated in the "measurement coordinate system" where the variables are R, A, E and their derivatives or in the "cartesian coordinate systems' in which the variables are x, y, z-position of the target and their derivatives In the latter case, an inertial reference frame is used.

Selecting a Coordinate System

It can be shown that in the Cartesian coordinate system, the dynamical equation for the target motion is linear whereas the measurement equation is non-linear in states. On the other hand, in the measurement coordinate system, the dynamical equation is non-linear whereas the measurement equation is linear in state variable. Therefore in both the coordinate systems, an extended Kalman filter may be used for state estimation. It is well known that an extended Kalman filter has a bias in its estimates [Jazwinski, 1970]. A detailed study on this issue was made by Mehra [1971]. He observed that this bias is related to the non-linearities in the equations of motion and the measurement system By choosing different coordinate systems, one can alter these non-linearities. In particular, if one uses the coordinate system in which the measurements are linear, the bias due to measurement non-linearity is eliminated This does not necessarily mean that the total bias is reduced since the bias from other non-linearities might increase Finally, he has shown that an extended Kalman filter that uses the measurement coordinate system has less bias and less rms error than a Cartesian extended Kalman filter that uses the Cartesian coordinate system This fact certainly favors the use of the range (R), azimuth (A), and elevation (E) angle variables in the SDI tracking problem formulation But Song and Speyer (1984) has shown that for a class of non-linear measurement equations known as 'modifiable functions," the structure of the optimal filter is linear. They have called this filter a "Modified Gain Extended Kalman Filter" (MGEKF). Fortunately, many "bearings only" measurement equations belong to the class of "modifiable functions" and we have found that we can profitably use this concept to solve the distributed tracking algorithm A

linear filter is certainly more attractive than a nonlinear filter. Motivated by these observations, we will use the Cartesian coordinate system throughout this project. In this coordinate system, the dynamic equation is linear whereas the measurement equation is nonlinear in states

1.3 Summary of Phase I Activities

The problem of distributed detection and estimation has lately received considerable attention in the literature Therefore, as a first task in Phase I, relevant publications were surveyed The useful results from these papers were collated and several modifications/exter ions were made that are appropriate for the problem at hand Some of the related works on which the proposed technique heavily relies are:

- Speyer (1979) considered the discrete-time distributed estimation problem assuming that all sensors communicated information to all other sensors in the network in a broadcast mode. By formulating a linear Gaussian measurement problem with a Gauss-Markov state equation, he was able to show that the optimum estimate at each sensor could be obtained using its own Kalman filter. Each sensor sends its local state estimate to each other sensor at each time step which then constructs the globally optimum state The expense of distributed estimation is that an additional data-dependent vector of dimension equal to the state vector dimension has to be communicated from each sensor to every other sensor at each time step. This solution is attractive since each sensor computes the optimal (minimum variance) state estimate and only a relatively small amount of communication is required between sensors.

- Willsky et. al. (1982) considered a more general case than Speyer (1979) and formulated the problem in continuous time domain They considered the case where the local models assumed by various sensors are possibly different than

the true global model They also solved the distributed smoothing problem. Their results also show that each node must send 2 vectors - its state estimate and a data-dependent vector to the central coordinator which constructs the globally optimum estimate.

- Castanon and Teneketzis (1985) have developed algorithms for distributed nonlinear estimation and showed that if each sensor sends a set of local sufficient statistics to each other sensor, each sensor can construct the global centralized distribution. These results assume that all sensors have the same a priori knowledge of system uncertainties, that all (two-way) communication links are unfailed, and that all communication channels are memoryless and independent

But the problem with correlated sensor noise was not addressed in these literature.

In Phase I, we have extended the results of Willsky et. al. (1982) to show that only one vector, instead of two, need to be sent to the global coordinator. We have also developed an expression that shows the effect of initial uncertainty of various sensors. Also, we have extended the result of Speyer (1979) to include the case where the local models are not necessarily identical to the global (true) model.

Three schemes have been developed for solving the problem with correlated sensor noise. The first two of these are straightforward application of the Kalman filter at a central node. In the first scheme, all the sensor data are transmitted to a central node and a Kalman filter is applied to the aggregated data. The second scheme also consists of the centralized processing but before the Kalman filter is applied, the sensor measurements are transformed into new measurements that are uncorrelated across the sensors. Clearly, there is no advantage of the second scheme over the first. However, the third scheme is truly a decentralized one where the observed data is processed locally and then the processed data

is sent to the central node where it is fused with the other incoming data into the globally optimal estimate.

We have fond that nonlinear measurement equations can be dealt with in the distributed tracking problem by utilizing the "modifiable fun tion" concept of Song and Spe, (1984). In this phase we have successfully applied this concept to the problem of "bearings only" measurements. In Phase II, this technique will be extended to the case of distributed processing

The remainder of this report is organized as follows: In Chapter 2, the distributed processing problem has been described in continuous time domain along the line of Willsky et. al (1982) and their results have been extended. In Chapter 3, the discrete time version of the tracking problem has been presented where the results of Speyer (1979) are extended to include the case of possibly dissimilar global and local models. It will be assumed in Chapter 2 and 3 that the sensors are uncorrelated. The case of correlated sensor noise will be dealt with in Chapter 4 where three techniques will be proposed for solving this problem The case of nonlinear measurement equations will be dealt with in Chapter 5. Finally, conclusions and future recommendations will be given in Chapter 6.

CHAPTER 2

DECENTRALIZED ESTIMATION: CONTINUOUS TIME CASE

2.1 Introduction

In an SDI environment a target is tracked by many sensors placed in orbiting satellites at desired altitudes. These satellites are linked via communication channels. The goal is to generate the optimal estimate of the state trajectory of the target from these observations. Clearly, the simplest scheme by which it can be done is to transmit at each instant of time all the sensor data to a central coordinator where the optimal trajectory can be computed conditional on the measurement history. Although this is a simple scheme, it suffers from many drawbacks. For example, the communication introduces a finite delay in the observed data and, in addition, adds to it noise from the communication channel. It may be recalled that the observed data was, to begin with, already corrupted from the sensor noise. The most serious criticism of this scheme is that all the computations are done at a central node and in the event of a failure (soft or hard) of this node, catastrophic failure of the system performance takes place This can be avoided if the sensor data is processed locally to obtain, say, locally optimum state estimate conditional on the observed data by that sensor only and then these estimates are sent to the centralized coordinator where the globally optimum estimate will be reconstructed This scheme is known as "decentralized" or "distributed" estimation. In this scheme, the system performance will suffer graceful degradation in the event of a sensor failure. Moreover, the computational burden is shared by all the sensors without imposing a heavy workload on the centralized coordinator. In order to develop redundancy into the system, each sensor should ideally act as a centralized coordinator in the sense that each sensor will receive the estimate from the other sensors and generate identical global estimates.

A difficult part of the decentralized scheme, however, is the development of a fusion scheme which combines in an appropriate fashion all the local estimates into a global estimate. The primary purpose of this chapter is to formulate this problem for the continuous time case, review the existing results and finally extend these results appropriately.

A practical combining (fusion) algorithm must take into account the communication delay and the channel noise, but it will make the analysis too complicated. Therefore in Phase I, we have considered an ideal situation where the communication delay and channel noise have been neglected These issues will be dealt with in Phase II In this chapter we have emphasized the development of a mixing algorithm under the assumption that all sensors are in a broadcast mode and receive information from the other sensors instantaneously. We have primarily followed the works of Willsky et al. (1982) in the development of this algorithm.

A summary of this chapter is as follows In Section 2.2, the decentralized estimation scheme is developed under the above simplifying assumptions. Willsky et al. (1982) has shown that in order to obtain the globally optimum state estimate, each node (sensor) must send to the central coordinator its own state estimate. But the structure of the local and central nodes is complicated. We have extended this result to simplify the computational complexity. In this section we have further assumed that the global model is not necessarily the same as the local model each sensor (agent) has possibly a different model about the target. The global model is the true model of the target. This assumption is relaxed in Section 2.3 where we have assumed that the global model is the same as the local model. This case was originally analyzed by Speyer (1979) It is shown in this case that the globally optimal estimate is a linear combination of the local estimate plus a dynamic

correction that results because of the distributed nature of the algorithm. Finally, some conclusions are drawn in Section 2.4

2.2 Decentralized Estimation and Transmission Requirements

As mentioned before, the problem of decentralized estimation will be formulated under the following assumptions:

- The dynamics are evolving in continuous time; the discrete time case will be analyzed in the next chapter.
- Only one target is to be tracked.
- The target dynamics are linear, but may possibly be time varying.
- The measurement equations are linear in states; the nonlinear equations will be treated in Chapter 5, using "modifiable function" concept.
- There is no communication delay
- There is no channel noise.
- The nodes are in a broadcast node, i.e. each node can communicate with all the other nodes. The ring structure of a realistic sensor configuration will be dealt with in Phase II.
- The sensors are uncorrelated from each other; the case of correlated sensors will be analyzed in Chapter 4.

First, let us fix the terminologies that will be used throughout this report. We assume a scenario where multiple sensors are observing an SDI target. We also assume that these sensors are equipped with the data processing capabilities. The sensors are also known as nodes, local agents or local processors. The target state model is the true model and its state estimate is to be computed from the sensor measurements. The true model is also known as the central or global model. In this section we consider a general case: we

fusion center, also known as a mixing mode where all the local estimates will be combined to obtain the global estimate. Since all the sensors are in a broadcast mode, each sensor receives raw observations or processed data from the other sensors. We assume here that each sensor acts as a fusion center and combines the data from the other sensors into a global estimate. Since each sensor is producing identical global estimates, this scheme introduces enough redundancy into the system and makes it robust against sensor failure.

Suppose that there are J sensors. The global and the local models are given in the following.

Global Model

$$\dot{x}(t) = A(t)x(t) + w(t)\,; \quad t \geq t_0\,; \tag{2.1a}$$

$$z_j(t) = C_j(t)x(t) + v_j(t)\,; \quad j = 1, 2, \ldots J \tag{2.1b}$$

Various statistics relating to the model are:

$$E\,[x(t_0)] = \bar{x}(t_0)\,; \quad E\,[(x(t_0) - \bar{x}(t_0))\,(x(t_0) - \bar{x}(t_0))^T] = \Sigma(t_0)\,;$$
$$E\,[w(t)] = 0\,; \quad E\,[w(t)w^T(\tau)] = Q(t)\delta(t-\tau)\,;$$
$$E\,[v_j(t)] = 0\,, \quad E\,[v_j(t)v_k^T(\tau)] = R_j(t)\delta(t-\tau)\delta_{jk}R_j(t) = R_j^T(t) > 0\,.$$

We assume that $w(t)$ is independent of $x(t_0)$ and of $v_j(t)$ for all $j = 1, 2, \ldots J$. It is also assumed that $v_j(t)$ is independent of $x(t_0)$.

Although the true model is given by (2.1), the sensors may not have the complete knowledge of it Suppose the j-th sensor is assuming the following model about the target: this model is also known as the local model.

Local Models

$$x_j(t) = A_j(t)x_j(t) + w_j(t) \tag{2.2a}$$

$$z_j(t) = H_j(t)x_j(t) + v_j(t)\ ; \quad j = 1, 2, \ldots J \tag{2.2b}$$

$$E\,[x_j(t_0)] = \bar{x}_j(t_0)\ ; \quad E\,[(x_j(t_0) - \bar{x}_j(t_0))\,(x_j(t_0) - \bar{x}_j(t_0))^T] = \Sigma_j(t_0)\ ;$$
$$E\,[w_j(t)] = 0\ ; \quad E\,[w_j(t)w_j^T(\tau)] = Q_j(t)\delta(t-\tau)\ ;$$
$$E\,[v_j(t)] = 0\ ; \quad E\,[v_j(t)v_k^T(\tau)] = R_j(t)\delta(t-\tau)\delta_{jk}\ ; \quad R_j(t) = R_j^T(t) > 0\ .$$

As before, we assume that $v_j(t)$ is independent of $x_j(t_0)$ and of $w_j(t)$ for all $j = 1, 2, \ldots J$. Notice that the measurement noise process has been assumed to be the same for both global and local models (2.1) and (2.2). This is a plausible assumption, because whatever the underlying model is, the sensor is always the same and therefore the sensor noise ought to be the same too. However, in order to make the problem well posed we need to impose the following additional relationship between the local and global models:

$$C_j(t) = H_j(t)M_j(t)\ ; \quad j = 1, 2, \ldots J \tag{2.3}$$

where $M_j(t)$ is a time varying matrix of appropriate dimension. Besides the requirement of (2.3) and that the sensor noise $v_j(t)$ are the same, the two models can be totally arbitrary.

General Solution Without Decentralized Considerations

The globally optimal estimate $\hat{x}(t)$ using all the information is given by

$$\hat{x}(t) = E\,[x(t)\ |\ z_1(\xi), z_2(\xi), \ldots z_J(\xi)\ ; \quad t_0 \le \xi \le t\,] \tag{2.4a}$$

where $E[\bullet]$ is the expectation operator and can be implemented at a central node if all the sensor data $z_j(t)$, $j = 1, 2, \ldots J$ are sent to this node. In order to evaluate $\hat{x}(t)$ first define

$$z(t) \stackrel{\Delta}{=} \begin{bmatrix} z_1(t) \\ z_2(t) \\ \cdot \\ \cdot \\ \cdot \\ z_J(t) \end{bmatrix}\ ; \quad C(t) \stackrel{\Delta}{=} \begin{bmatrix} C_1(t) \\ C_2(t) \\ \cdot \\ \cdot \\ \cdot \\ C_J(t) \end{bmatrix}\ ; \quad v(t) \stackrel{\Delta}{=} \begin{bmatrix} v_1(t) \\ v_2(t) \\ \cdot \\ \cdot \\ \cdot \\ v_J(t) \end{bmatrix}\ ; \tag{2.4b}$$

then all the sensor data can be compactly represented as

$$z(t) = C(t)x(t) + v(t) . \tag{2.4c}$$

The optimal estimate $\hat{x}(t)$ conditional upon the measurement history is then

$$\hat{x}(t) = E\,[x(t) \mid z(\xi) , \quad t_0 \le \xi \le t\,]$$

which is propagated as

$$\dot{\hat{x}}(t) = A(t)\hat{x}(t) + K^c(t)\upsilon^c(t) , \tag{2.5a}$$

$$\upsilon^c(t) = z(t) - C(t)\hat{x}(t) , \tag{2.5b}$$

$$K^c(t) = P(t)C^T(t)R^{-1}(t) . \tag{2.5c}$$

Here $K^c(t)$ is the centralized Kalman gain. The covariance matrix $P(t)$ is the solution of the Riccati equation

$$\begin{aligned} \dot{P}(t) &= A(t)P(t) + P(t)A^T(t) + Q(t) - P(t)C^T(t)R^{-1}(t)C(t)P(t) . \\ P(t_0) &= \Sigma(t_0) . \end{aligned} \tag{2.5d}$$

$R(t)$ is the covariance matrix associated with the sensor noise $v(t)$, i.e.

$$E\,[v(t)v^T(\tau)] = R(t)\delta(t-\tau) \tag{2.6a}$$

The complexity of the decentralized algorithm depends upon the structure of the $R(t)$ matrix. Since we are considering the case of uncorrelated sensor noise, various subvectors of $v(t)$ of (2.4b) are uncorrelated. Therefore

$$R(t) = \text{block diag}\,\{R_j(t)\} . \tag{2.6b}$$

R(t) will be a "full" matrix if the sensors are correlated. However using the block structure of R(t) matrix the centralized Kalman filter of Equation (2.5) for the uncorrelated sensor noise can be written as:

$$\dot{\hat{x}}(t) = A(t)\hat{x}(t) + \sum_{j=1}^{J} K_j^c(t)\upsilon_j^c(t) \,, \tag{2.7a}$$

$$\upsilon_j^c(t) = z_j(t) - c_j(t)\hat{x}(t) \,, \tag{2.7b}$$

$$K_j^c(t) = P(t)C_j^T(t)R_j^{-1}(t) \,, \tag{2.7c}$$

$$\dot{P}(t) = A(t)P(t) + P(t)A^T(t) + Q(t) - \sum_{j=1}^{J} P(t)C_j^T(t)R_j^{-1}(t)C_j(t)P(t) \,. \tag{2.7d}$$

Equation (2.7a) can further be simplified to

$$\dot{\hat{x}}(t) = F(t)\hat{x}(t) + \sum_{j=1}^{J} K_j^c(t)z_j(t) \tag{2.8a}$$

where

$$F(t) = A(t) - \sum_{j=1}^{J} P(t)C_j^T(t)R_j^{-1}(t)C_j(t) \,. \tag{2.8b}$$

An implementation of Equation (2.8) is shown in the block diagram in Figure 2 1. Notice that the gain matrices $K_j^c(t)$, j = 1, 2, ...J are precomputable and can be stored apriori.

Decentralized Estimation Using Local Data Processing

Now consider the local models of Equation (2.2). The optimal estimate $\hat{x}_j(t)$ at the j-th node (sensor) is constructed only from the local observation $z_j(t)$ from that sensor. Therefore

$$\hat{x}_j(t) = E\,[x_j(t) \mid z_j(\xi)\,, \quad t_0 \le \xi \le t]$$

which evolves as

$$\dot{\hat{x}}_j(t) = A_j(t)\hat{x}_j(t) + K_j^d(t)\upsilon_j^d(t)\,, \tag{2.9a}$$
$$\upsilon_j^d(t) = z_j(t) = H_j(t)\hat{x}_j(t)\,,$$
$$K_j^d(t) = P_j(t)H_j^T(t)R_j^{-1}(t)\,.$$

Here $K_j^d(t)$ is the decentralized Kalman gain at the j-th node. The error covariance matrix $P_j(t)$ can be computed from either of the following forms of the Riccati equation:

$$\dot{P}_j(t) = A_j(t)P_j(t) + P_j(t)A_j^T(t) + Q_j(t) - P_j(t)C_j^T(t)R_j^{-1}(t)C_j(t)P_j(t) \tag{2.9b}$$

$$\frac{d}{dt}(P_j^{-1}(t)) = -P_j^{-1}(t)A_j(t) - A_j^T(t)P_j^{-1}(t) - P_j^{-1}(t)Q_j(t)P_j^{-1}(t)$$
$$+ H_j^T(t)R_j^{-1}(t)H_j(t) \tag{2.9c}$$

with the initial condition

$$P_j(t_0) = \Sigma_j(t_0)$$

Equation (2.9a) further reduces to

$$\dot{\hat{x}}_j(t) = F_j(t)\hat{x}_j(t) + K_j^d(t)z_j(t) \tag{2.10a}$$

where

$$F_j(t) = A_j(t) - P_j(t)H_j^T(t)R_j^{-1}(t)H_j(t) = A_j(t) - K_j^d(t)H_j(t)\,. \tag{2.10b}$$

The local estimator for the j-th node is shown in Figure 2.2.

Relationship Between the Local and Global Estimates

To find the connection between the two estimates, we will use the relationship $C_j(t) = H_j(t)M_j(t)$ between the local and global models. This relationship has been described earlier in (2.3) However using this relationship, the global estimate of (2.8a) can be written as

$$\dot{\hat{x}}(t) = F(t)\hat{x}(t) + \sum_{j=1}^{J} P(t)M_j^T(t)H_j^T(t)R_j^{-1}(t)z_j(t)\,. \tag{2.11a}$$

From (2.10) we also have

$$P_j(t)H_j^T(t)R_j^{-1}(t)z_j(t) = \dot{\hat{x}}_j(t) - F_j(t)\hat{x}_j(t)\,. \tag{2.11b}$$

Using (2.11b) in (2.11a), we obtain as the global estimate

$$\dot{\hat{x}}(t) = F(t)\hat{x}(t) + \sum_{j=1}^{J} G_j(t)\,[\dot{\hat{x}}(t) - F_j(t)\hat{x}_j(t)]\,, \tag{2.12a}$$

where

$$G_j(t) = P(t)M_j^T(t)P_j^{-1}(t)\,. \tag{2.12b}$$

By straightforward manipulation, it can be shown that

$$\hat{x}(t) = \phi_F(t, t_0)\left[\hat{x}(t_0) - \sum_{j=1}^{J} G_j(t_0)\hat{x}_j(t_0)\right] + \sum_{j=1}^{J} G_j(t)\hat{x}_j(t)$$
$$+ \sum_{j=1}^{J} \int_{t_0}^{t} \phi_F(t, \tau)K_j(\tau)\hat{x}_j(\tau)d\tau\,, \tag{2.13a}$$

where $\phi_F(t, \tau)$ is the transition matrix associated with the dynamics matrix $F(t)$ and

$$
\begin{aligned}
K_j(t) &= F(t)G_j(t) - \dot{G}_j(t) - G_j(t)F_j(t) \\
&= [P(t)M_j^T(t)P_j^{-1}(t)Q_j(t)P_j^{-1}(t) - Q(t)M_j^T(t)P_j^{-1}(t)] \\
&\quad + [P(t)M_j^T(t)A_j^T(t)P_j^{-1}(t) - P(t)A^T(t)M_j^T(t)P_j^{-1}(t) \\
&\quad - P(t)\dot{M}_j^T(t)P_j^{-1}(t)]
\end{aligned}
\tag{2.13b}
$$

The indicated inverse of the matrices are assumed to exist. However this is the key equation relating the local and global estimates. This equation also shows the effect of initial conditions at various sensors on the global estimate. It is obvious from (2.13a) that $\hat{x}(t)$ can also be generated dynamically from the following equations:

$$
\dot{\xi}(t) = F(t)\xi(t) + \sum_{j=1}^{J} K_j(t)\hat{x}_j(t)\,; \quad \xi(t_0) = \hat{x}(t_0) - \sum_{j=1}^{J} G_j(t_0)\hat{x}_j(t_0) \tag{2.14a}
$$

$$
\hat{x}(t) = \xi(t) + \sum_{j=1}^{J} G_j(t)\hat{x}_j(t) \tag{2.14b}
$$

These equations were first established by Willsky et al. (1982). Equations (2.14) clearly reveal the structure of the decentralized estimation: the measurements at the j-th sensor is processed locally to obtain the optimal estimate $\hat{x}_j(t)$ which is then sent to the centralized coordinator, the coordinator then constructs $\hat{x}(t)$ dynamically according to (2 14). Notice that all the matrices in (2.14) are precomputable and can be stored apriori. The block diagram representing these equations are shown in Figure 2.3.

It can be seen from Figure 2.3 that there is a considerable amount of computational burden at the central node and the structure of the combining algorithm is rather complex The computational burden and the complexity of the centralized coordinator can be reduced as follows. Define $h_j(t)$ as a data dependent vector for the j-th node which evolves as

$$
\dot{h}_j(t) = F(t)h_j(t) + K_j(t)\hat{x}_j(t)\,, \tag{2.15a}
$$

with initial condition

$$h_j(t_0) = -G_j(t_0)\hat{x}_j(t_0) . \tag{2.15b}$$

The global estimate then simplifies to

$$\hat{x}(t) = \phi_F(t, t_0)\hat{x}(t_0) + \sum_{j=1}^{J}[G_j(t)\hat{x}_j(t) + h_j(t)] . \tag{2.15c}$$

This form was first obtained by Speyer (1979) for a simplified case when the global model is the same as the local model We will discuss this case in the next section. The implementation of this form is shown in Figure 2.4. In this scheme, each node j must send two vectors - $\hat{x}_j(t)$ and $h_j(t)$ to the central node which then constructs the global estimate $\hat{x}(t)$. Notice also that the computational burden of the central node has been reduced considerably but only at the expense of the additional complexity and burden at the local nodes. By simple block diagram manipulation of Figure 2.4, the complexity of the central node can be reduced further. To see this mathematically, define a node-dependent dynamic vector $\eta_j(t)$ which is generated as:

$$s_j(t) = F(t)s_j(t) + K_j(t)\hat{x}_j(t) ; \quad s_j(t_0) = -G_j(t_0)\hat{x}_j(t_0) \tag{2.16a}$$

$$\eta_j(t) = s_j(t) + G_j(t)\hat{x}_j(t) \tag{2.16b}$$

Clearly, then the global estimate is given by

$$\hat{x}(t) = \phi_F(t, t_0)\hat{x}(t_0) + \sum_{j=1}^{J}\eta_j(t) \tag{2.16c}$$

which can be implemented as in Figure 2.5. Notice that in this scheme, only one vector $\eta_j(t)$ is to be sent to the central node in contrast to 2 vectors in the earlier scheme of (2.15). Thus, the transmission requirement has been reduced by 50%. If $\hat{x}(t_0) = 0$, then the central

node is simply an adder, it merely adds the incoming vectors from the various nodes. Notice also that the complexity and the computational burden of the central node has been reduced to minimum but only by burdening each of the sensor nodes with these problems.

Equation (2.15) can be manipulated further to reduce the complexity of the local processors as follows Define

$$q_j(t) = G_j(t)\hat{x}_j(t) + h_j(t) \tag{2.17a}$$

where the evolution of $h_j(t)$ with its initial condition is given in Equations (2.15a) and (2.15b) respectively. Since $h_j(t_0) = -G_j(t_0)\hat{x}_j(t_0)$, the initial condition of $q_j(t)$ is

$$q_j(t_0) = 0 \tag{2.17b}$$

The global estimate can be written in terms of $q_j(t)$ using (2.15c):

$$\hat{x}(t) = \phi_F(t, t_0)\hat{x}(t_0) + \sum_{j=1}^{J} q_j(t) . \tag{2.17c}$$

A recursion for $q_j(t)$ can be found as follows: Differentiating both sides of (2.17a) gives

$$\dot{q}_j(t) = \dot{G}_j(t)\hat{x}_j(t) + G_j(t)\dot{\hat{x}}(t) + \dot{h}_j(t) . \tag{2.17d}$$

Using Equations (2.9a) and (2.15a) in the above expression, we get

$$\begin{aligned} q_j(t) = [\dot{G}_j(t) + G_j(t)F_j(t) + K_j(t)]\hat{x}_j(t) + G_j(t)P_j(t)H_j^T(t)R_j^{-1}(t)z_j(t) \\ + F(t)h_j(t) . \end{aligned} \tag{2.17e}$$

Finally, using expression for $K_j(t)$ from (2.13b) in (2.17e) and simplifying, we have,

$$\dot{q}_j(t) = F(t)q_j(t) + \tilde{K}_j(t)z_j(t) \tag{2.17f}$$

where

$$\bar{K}_j(t) = P(t)M_j^T(t)H_j^T(t)R_j^{-1}(t) .$$

The decentralized estimation scheme given in (2.17) above is shown Figure 2.6. This scheme may be compared with the one in Figure 2.5. Notice that the complexity of the local processor has been reduced and yet the transmission requirement has remained the same each local processor need to send to the mixing node only one data dependent vector $q_j(t)$ of the dimension of the true (global) state.

An important remark is in order here. The fact that the sensors are uncorrelated has played a crucial role in developing the decentralized scheme here. It is the block diagonal structure of the R(t) matrix which has enabled us to decompose the centralized estimation scheme of (2.5) into additive form of (2.7). The centralized Kalman gain $K^c(t)$ can be partitioned as $K^c(t) = [K_1^c(t), K_2^c(t) \ldots K_J^c(t)]$ each block of which corresponds to a sensor. When the sensors are correlated, i.e. the R(t) matrix is of "full" structure, it is not clear how to decompose the global estimation problem into various subproblems.

2.3 The Special Case of Identical Local and Global Models

In this section we consider a special case of the more general formulation of the last section. Specifically, we assume that the models used by the local processors are identical to the global model. This case was first examined by Speyer (1979). Under this assumption

$$A_j(t) = A(t)\ ,\ \ Q_j(t) = Q(t)\ ,\ \ C_j(t) = H_j(t)\ ,\ \ M_j(t) = I \quad \text{for all j.} \qquad (2.17)$$

In this case, the expression for $K_j(t)$ of Equation (2.13b) simplifies to

$$
\begin{aligned}
K_j(t) &= P(t)P_j^{-1}(t)Q(t)P_j^{-1}(t) - Q(t)P_j^{-1}(t) \\
&= [P(t)P_j^{-1}(t) - I]\, Q(t)P_j^{-1}(t)\,.
\end{aligned}
\tag{2.18}
$$

Since $M_j(t) = I$, we have set $\dot{M}_j(t) = 0$ to find the above expression. The other relevant variables for this case are

$$
\begin{aligned}
F(t) &= A(t) - \sum_{j=1}^{J} P(t)C_j^T(t)R_j^{-1}(t)C_j(t)\,; \\
F_j(t) &= A(t) - P_j(t)C_j^T(t)R_j^{-1}(t)C_j(t)\,; \\
G_j(t) &= P(t)P_j^{-1}(t)\,; \\
K_j^c(t) &= P(t)C_j^T(t)R_j^{-1}(t)\,; \\
K_j^d(t) &= P_j(t)C_j^T(t)R_j^{-1}(t)\,.
\end{aligned}
\tag{2.18b}
$$

An important consequence of this simplification is that the global estimate of (2.14b) can be written as

$$
\hat{x}(t) = \xi(t) + P(t)\sum_{j=1}^{J} P_j^{-1}(t)\hat{x}_j(t)
\tag{2.19}
$$

Note that the second term in the expression for $\hat{x}(t)$ is the usual expression for combining independent estimates However $\hat{x}_j(t)$ are not independent in general, and $\xi(t)$ represents a dynamic correction for this correlation.

2.4 Summary

In this chapter, we have formulated and solved the decentralized estimation problem for the continuous time case. We have primarily followed the work of Willsky et al. (1982) to derive the results of this chapter. The assumption that the sensors are uncorrelated has played a crucial role in developing these algorithms.

The multisensor data processing problem can be solved in two ways. All the sensor data can be transmitted to a central node and can be processed for globally optimum estimates. Alternatively, each sensor data can be processed locally for a locally optimum estimate which then can be transmitted to the central node. The central node combines all the local estimates to obtain the global estimate. The later scheme is the decentralized estimation and is preferable from the viewpoint of system survivability. In Section 2.2, a generalized decentralized estimation scheme has been formulated. Here we have assumed that each of the local processors may possibly have a different model which again is not necessarily the same as the global model. Several possible decentralized schemes have been presented in Figures 2.3 - 2.5. In the scheme of Figure 2.3, the local estimates are sent to the mixing node which in turn produces the globally optimum estimate dynamically. The structure of the mixing node is complex in this case. The complexity of this node can be reduced by transferring some of its computational burden to the local nodes as shown in Figure 2.4. But in this case, each local node must send 2 vectors to the mixing node. The transmission requirement of this scheme can be reduced to one vector by adding additional computations at the local nodes as shown in Figure 2.5. The results of this section show that the complexity and the computational burden of the central node can be traded by those of the local nodes. Finally, in Figure 2.6, the complexity of the local processor has been reduced without sacrificing the transmission efficiency.

In Section 2.3, a specialized case has been dealt with where the local models are assumed to be the same as the a global model. This case was first analyzed by Speyer (1979). Under this assumption, considerable simplification occurs in the algorithms and in the structure of the local and central processors.

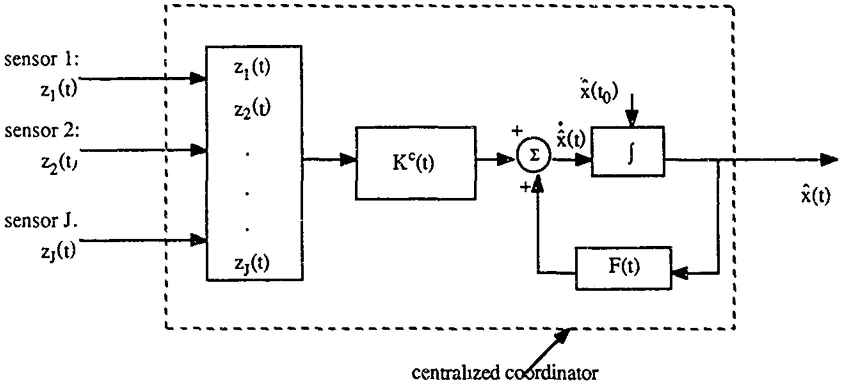
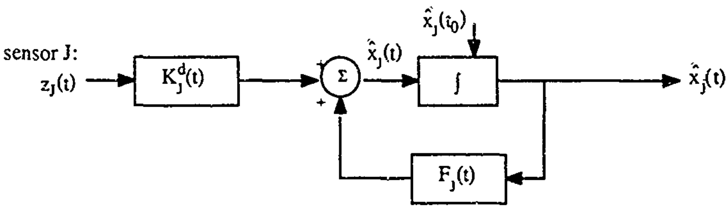


Figure 2.1: Centralized Estimation



2.2: Local Estimator: j-th Sensor

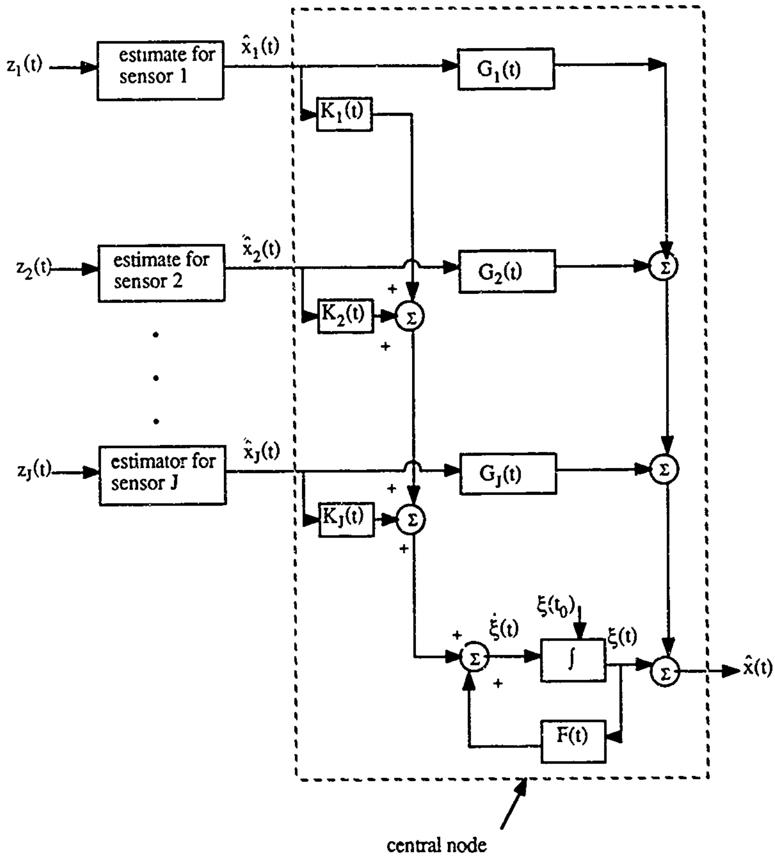


Figure 2.3: Decentralized Estimation: Basic Structure

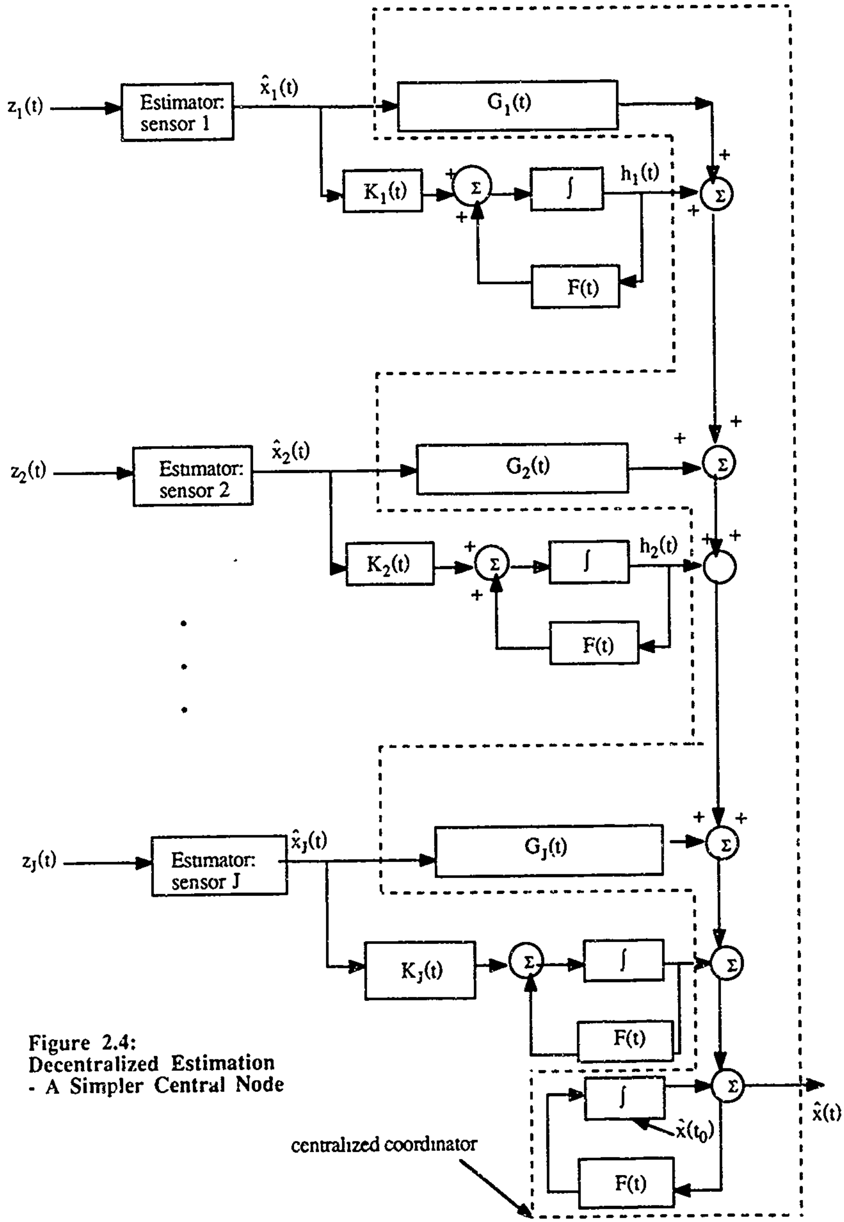


Figure 2.4:
Decentralized Estimation
- A Simpler Central Node

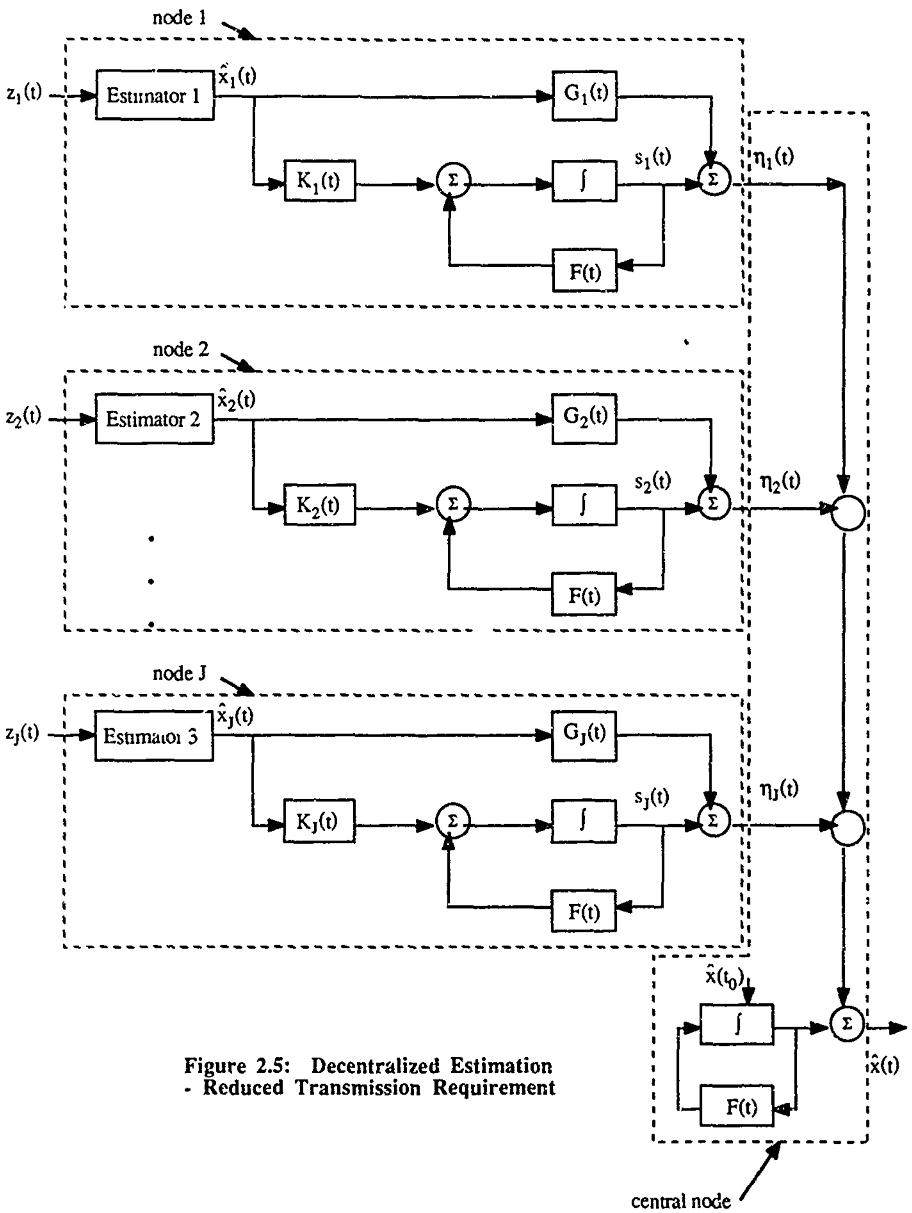


Figure 2.5: Decentralized Estimation - Reduced Transmission Requirement

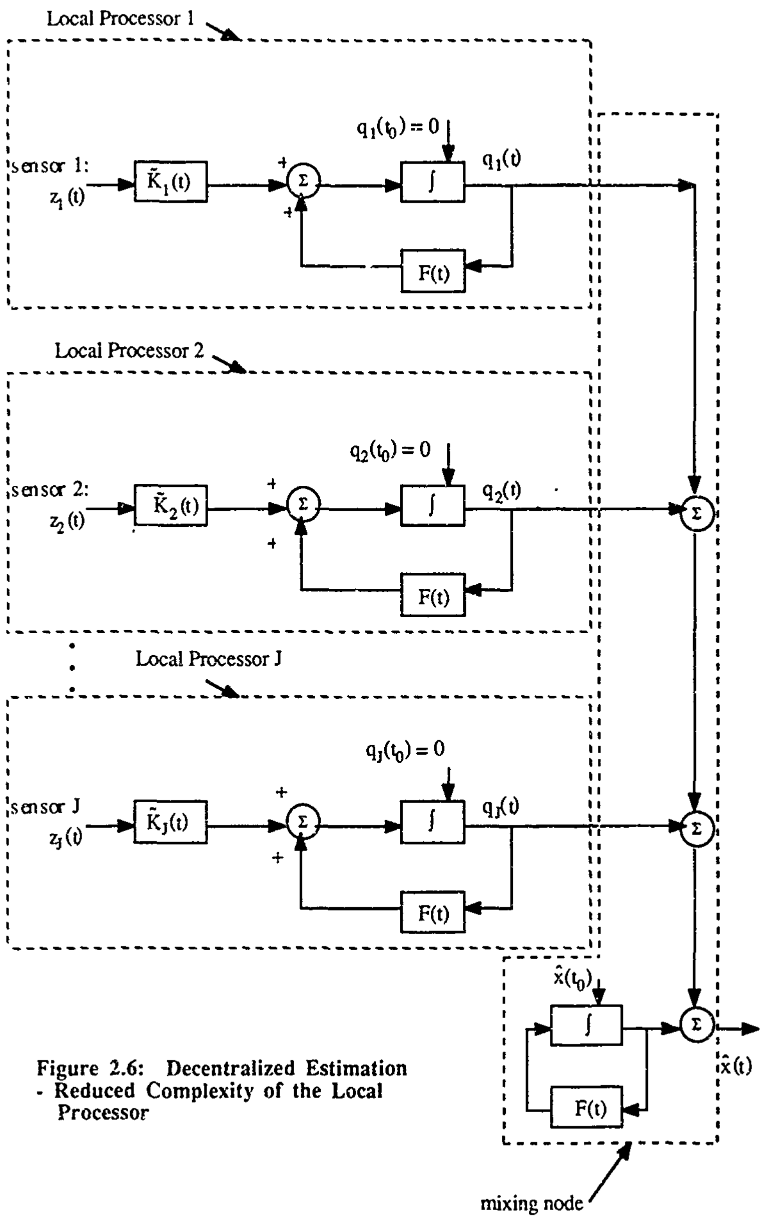


Figure 2.6: Decentralized Estimation - Reduced Complexity of the Local Processor

CHAPTER 3

DECENTRALIZED ESTIMATION: DISCRETE TIME VERSION

3.1 Introduction

The decentralized estimation problem for the continuous time case was analyzed in the last chapter In this chapter, the corresponding problem for the discrete time case will be presented. This problem was first analyzed by Speyer (1979), but for the special case of identical global and local models. In this chapter, we intend to extend his results to a more general case of non-identical global and local models. Some of these results have already been submitted before in various progress reports.

The motivation for the decentralized estimation problem and various underlying assumptions behind the problem formulation have already been presented in Chapter 2 and will not be repeated here. This chapter will be a brief one; we shall only present the corresponding results of the Chapter 2 following the same sequence of that chapter. The purpose of this chapter is twofold: first, for the sake of completeness of this report and second, to expose some implementation details that are distinct from the continuous time case. Explanations and interpretation of various results are same as in Chapter 2.

This chapter is organized as follows. In Section 3.2 the decentralized estimation problem will be formulated and the computational and transmission requirement of each local processor will be established. As in the last chapter, we assume that each local processor has two alternatives: either it can send the raw observation data to a central (mixing) node where the globally optimum estimate will be computed or it can process the raw data for the locally optimum (or some kind of) estimate which are then transmitted to a mixing node. The local estimates are then fused at the mixing node for the global estimate.

We assume that the local processing is always a superior alternative than the centralized processing. The reasons are given in Chapter 2. In this section, we shall present several schemes for local processing corresponding to those in Chapter 2. In Section 3.3, we shall simplify the derivations of Section 3.2 for the special case of identical and global models - this case was analyzed by Speyer (1979). Finally, some conclusions are provided in Section 3.4.

3.2 Decentralized Estimation and Transmission Requirements

As in the last chapter, we assume a scenario where multiple sensors are tracking a single SDI target. These sensors will also be referred to as nodes, local agents or local processors. The target model is called the global or central model and the goal is to estimate the target state from all the sensor measurements. In this section, we assume that the agents may not possibly have the full knowledge of the global model. Instead, these processors will have their own model (also known as local model) about the target state which they will use to estimate the target states. These estimates are then transmitted to a fusion center, also known as the mixing node, where all the local estimates are combined to obtain the global estimate.

Suppose there are J sensors. The global and local models are as follows.

Global Model

$$x(i+1) = A(i)\,x(i) + w(i)\,; \quad i \geq i_0 \tag{3.1a}$$

$$z_j(i) = C_j(i)\,x(i) + v_j(i)\,; \quad j = 1,2,\ldots J \tag{3.1b}$$

Various statistics about this model are as follows:

$$E\,[x(i_0)] = \bar{x}(i_0)\,;\ E\,[(x(i_0) - \bar{x}(i_0))\,(x(i_0) - \bar{x}(i_0))^T] = \Sigma(i_0)$$

$$E[w(i)] = 0; \quad E[w(i)w^T(m)] = Q(i)\delta_{im}$$

$$E[v_j(i)] = 0; \quad E[v_j(i)v_k^T(m)] = R_j(i)\delta_{im}\delta_{jk}; \; R_j(i) = R_j^T(i) > 0 .$$

We also assume that $w(i)$ is independent of $x(i_0)$ and of $v_j(i)$ for all $j=1,2,...J$ and $v_j(i)$ is independent of $x(i_0)$.

Local Models

$$x_j(i+1) = A_j(i)x_j(i) + w_j(i) \tag{3.2a}$$

$$z_j(i) = H_j(i)x_j(i) + v_j(i); \quad j=1,2,...J \tag{3.2b}$$

$$E[x_j(i_0)] = \bar{x}_j(i_0); \quad E[(x_j(i_0) - \bar{x}_j(i_0))(x_j(i_0) - \bar{x}_j(i_0))^T] = \Sigma_j(i_0)$$

$$E[w_j(i)] = 0; \quad E[w_j(i)w_k^T(m)] = Q_j(i)\delta_{im}\delta_{jk}$$

$$E[v_j(i)] = 0; \quad E[v_j(i)v_k^T(m)] = R_j(i)\delta_{im}\delta_{jk}$$

As before, we assume that $v_j(i)$ is independent of $x_j(i_0)$ and of $w_j(i)$ for all j and i. Further, for the well posedness of the problem, we assume the existence of a sequence of matrices $\{M_j(i)\}$ such that

$$C_j(i) = H_j(i)M_j(i); \quad j=1,2,...J \tag{3.3}$$

General Solution Without Decentralized Considerations

The globally optimal filtered state using all the information is given by

$$\hat{x}(i \mid i) = E[x(i) \mid z_1(\zeta), z_2(\zeta), ...z_J(\zeta); \quad i_0 \le \zeta \le i] \tag{3.4a}$$

which can be mechanized by a central node if it receives all the sensor data. In order to evaluate $\hat{x}(i \mid i)$, define

$$z(i) = \begin{bmatrix} z_1(i) \\ z_2(i) \\ \cdot \\ \cdot \\ \cdot \\ z_J(i) \end{bmatrix}, \quad C(i) = \begin{bmatrix} C_1(i) \\ C_2(i) \\ \\ \cdot \\ \cdot \\ C_J(i) \end{bmatrix}; \quad v(i) = \begin{bmatrix} v_1(i) \\ v_2(i) \\ \cdot \\ \cdot \\ \cdot \\ v_J(i) \end{bmatrix}. \tag{3.4b}$$

Then all the sensor data can be compactly written as

$$z(i) = C(i)x(i) + v(i). \tag{3.4c}$$

Therefore, the optimal filtered state conditional on the measurement history is

$$\hat{x}(i \mid i) = E[x(i) \mid z(\zeta) ; i_0 \le \zeta \le i]$$

which evolves as [Bryson and Ho (1975)]

$$\hat{x}(i \mid i) = \hat{x}(i \mid i-1) + K^c(i)\upsilon^c(i) ; \tag{3.5a}$$

$$\upsilon^c(i) = z(i) - C(i)\hat{x}(i \mid i-1) ; \tag{3.5b}$$

$$K^c(i) = P(i \mid i)C^T(i)R^{-1}(i) , \tag{3.5c}$$

$$\hat{x}(i+1 \mid i) = A(i)\hat{x}(i \mid i) ;$$

where $R(i)$ is the covariance matrix associated with $v(i)$ that has been defined in (3.4b) and $K^c(i)$ is the centralized Kalman gain. The covariance matrix $P(i \mid i)$ is the solution of the Riccati equation and follows from the following recursion:

$$P^{-1}(i \mid i) = P^{-1}(i \mid i-1) + C^T(i)R^{-1}(i)C(i) ; \tag{3.5d}$$

$$P(i+1 \mid i) = A(i)P(i \mid i)A^T(i) + Q(i) ; \tag{3.5e}$$

$$P(i_0 \mid i_0 - 1) = \Sigma(i_0) .$$

$R(i)$ is defined as

$$E\{v(i)v^T(m)\} = R(i)\delta_{i\,m}. \tag{3.6a}$$

Since we are considering the case of uncorrelated sensor noise,

$$R(i) = \text{block diag}\,\{R_j(i)\} \tag{3.6b}$$

where $R_j(i)$ is defined in (3.1) and (3.2) above.

(3.5a) can be simplified further as follows:

$$\hat{x}(i+1|i+1) = \hat{x}(i+1|i) + K^c(i+1)[z(i+1) - C(i+1)\hat{x}(i+1|i)] \tag{3.7a}$$

$$= F(i+1)\hat{x}(i|i) + K^c(i+1)z(i+1) \tag{3.7b}$$

where

$$F(i) = \tilde{F}(i)A(i-1)\,; \tag{3.7c}$$

$$\tilde{F}(i) = I - \sum_{j=1}^{J} P(i|i)C_j^T(i)R_j^{-1}(i)C_j(i)\,; \tag{3.7d}$$

and $K^c(i)$ is defined in (3.5c). A block diagram of this centralized scheme is shown in Figure 3.1.

Using the block diagonal structure of the R(i) matrix, (3.7a) simplifies to

$$\hat{x}(i+1|i+1) = F(i+1)\hat{x}(i|i) + \sum_{j=1}^{J} K_j^c(i+1)z_j(i+1) \tag{3.8a}$$

where

$$K_j^c(i) = P(i|i)C_j^T(i)R_j^{-1}(i)\,; \tag{3.8b}$$

$$P^{-1}(i|i) = P^{-1}(i|i-1) + \sum_{j=1}^{J} C_j(i)R_j^{-1}(i)C_j(i) . \tag{3.8c}$$

Notice that using (3.8c), $\tilde{F}(i)$ of (3.7d) can be written as

$$\tilde{F}(i) = P(i|i)P^{-1}(i|i-1) . \tag{3.8d}$$

Decentralized Estimation Using Local Data Processing

Now consider the local model of Equation (3.2). The local estimate at the j-th node is given by

$$\hat{x}_j(i|i) = E[x_j(i) \mid z_j(\zeta) ; \ i_0 \le \zeta \le i] ; \tag{3.9a}$$

which is propagated as

$$\hat{x}_j(i|i) = \hat{x}_j(i|i-1) + K_j^d(i)\upsilon_j^d(i) ; \tag{3.9b}$$

$$\upsilon_j^d(i) = z_j(i) - H_j(i)\hat{x}_j(i|i-1) ;$$

$$K_j^d(i) = P_j(i|i)H_j^T(i)R_j^{-1}(i) ;$$

$$\hat{x}_j(i+1|i) = A_j(i)\hat{x}_j(i|i) .$$

Here $K_j^d(i)$ is the decentralized Kalman gain associated with the j-th sensor. The error covariance matrix $P_j(i|i)$ follows the recursion

$$P_j^{-1}(i|i) = P_j^{-1}(i|i-1) + H_j^T(i)R_j^{-1}(i)H_j(i) ; \tag{3.9c}$$

$$P_j(i+1|i) = A_j(i)P_j(i|i)A_j^T(i) + Q_j(i) ;$$

$$P_j(i_0|i_0-1) = \Sigma_j(i_0) .$$

(3.9b) can be simplified further:

$$\hat{x}_j(i+1|i+1) = F_j(i+1)\hat{x}_j(i|i) + K_j^d(i+1)z_j(i+1) \qquad (3.10a)$$

where

$$F_j(i) + \tilde{F}_j(i)A_j(i-1)$$

$$\tilde{F}_j(i) = I - P_j(i|i)H_j^T(i)R_j^{-1}(i)H_j(i)$$

$$= P_j(i|i)P_j^{-1}(i|i-1) \qquad (3.10b)$$

and the last equality has been derived from (3.9c). The local estimator for the j-th node is shown in Figure 3.2.

Relationship Between the Global and Local Estimates

Using the relationship $C_j(i)=H_j(i)M_j(i)$ in (3.8a), we have the global estimate as

$$\hat{x}(i+1|i+1) = F(i+1)\hat{x}(i|i)$$

$$+\sum_{j=1}^{J} P(i+1|i+1)M_j^T(i+1)H_j^T(i+1)R_j^{-1}(i+1)z_j(i+1) \qquad (3.11b)$$

Rearranging (3.10a) after substituting the value of $K_j^d(i+1)$ gives

$$H_j^T(i+1)R_j^{-1}(i+1)z_j(i+1) = P_j^{-1}(i+1|i+1)[\hat{x}_j(i+1|i+1)$$

$$-F_j(i+1)\hat{x}_j(i|i)] \qquad (3.11b)$$

Now, using (3.11b) in (3.11a), we get

$$\hat{x}(i+1|i+1) = F(i+1)\hat{x}(i|i) + \sum_{j=1}^{J} P(i+1|i+1)\, M_j^T(i+1)$$

$$\times P_j^{-1}(i+1|i+1)[\hat{x}_j(i+1|i+1) - F_j(i+1)\hat{x}_j(i|i)) \qquad (3.11c)$$

Since $\hat{x}_j(i|i) = A_j^{-1}(i)\hat{x}_j(i+1|i)$, we have

$$
\begin{aligned}
F_j(i+1)\hat{x}_j(i \mid i) &= F_j(i+1)A_j^{-1}(i)\hat{x}_j(i+1 \mid i) \\
&= \tilde{F}_j(i+1)\hat{x}_j(i+1 \mid i) \\
&= P_j(i+1 \mid i+1)P_j^{-1}(i+1 \mid i)\hat{x}_j(i+1 \mid i)\,. \qquad (3.12a)
\end{aligned}
$$

Using (3.12a) in (3.11c) gives the global estimate as

$$
\begin{aligned}
\hat{x}(i+1 \mid i+1) = F(i+1)\hat{x}(i \mid i) + \sum_{j=1}^{J} P(i+1 \mid i+1)M_j^T(i+1)[P_j^{-1}(i+1 \mid i+1) \\
\times \hat{x}_j(i+1 \mid i+1) - P_j^{-1}(i+1 \mid i)\hat{x}_j(i+1 \mid i)]\,. \qquad (3.12b)
\end{aligned}
$$

This expression suggests that $\hat{x}(i \mid i)$ is of the form

$$
\hat{x}(i \mid i) = \sum_{j=1}^{J} [P(i \mid i)M_j^T(i)P_j^{-1}(i \mid i)\hat{x}_j(i \mid i) + h_j(i)] \qquad (3.13a)
$$

where $h_j(i)$ is a node-dependent vector of the dimension of the state. This implies

$$
\begin{aligned}
\hat{x}(i+1 \mid i+1) = \sum_{j=1}^{J} [P(i+1 \mid i+1)M_j^T(i+1)P_j^{-1}(i+1 \mid i+1)\hat{x}_j(i+1 \mid i+1) \\
+ h_j(i+1)] \qquad (3.13b)
\end{aligned}
$$

To find a recursion for $h_j(i+1)$, (3.13a) is substituted in (3.12b) for $\hat{x}(i \mid i)$ and then comparing it with (3.13b), we obtain

$$
\begin{aligned}
h_j(i+1) = F(i+1)h_j(i) + F(i+1)P(i \mid i)M_j^T(i)P_j^{-1}(i \mid i)\hat{x}_j(i \mid i) \\
- P(i+1 \mid i+1)M_j^T(i+1)P_j^{-1}(i+1 \mid i)\hat{x}_j(i+1 \mid i) \qquad (3.13c)
\end{aligned}
$$

Since $\hat{x}_j(i \mid i) = A_j^{-1}(i)\hat{x}_j(i+1 \mid i)$, the above equation can be compactly written as:

$$
h_j(i+1) = F(i+1)h_j(i) + G_j(i+1)\hat{x}_j(i+1 \mid i) \qquad (3.14a)
$$

where

$$F(i+1) = \tilde{F}(i+1)A(i),$$
$$= P(i+1|i+1)P^{-1}(i+1|i)A(i); \tag{3.14b}$$

$$G(i+1) = P(i+1|i+1)P^{-1}(i+1|i)A(i)P(i|i)M_j^T(i)P_j^{-1}(i|i)$$
$$\times A_j^{-1}(i) - P(i+1|i+1)M_j^T(i+1)P_j^{-1}(i+1|i). \tag{3.14c}$$

We have used (3.7c) and (3.8d) in establishing (3.14b) which in turn has been used to find (3.14c). We have assumed that the inverse for the appropriate matrices exists in the above expression The initial condition for $h_j(i)$ can be evaluated by solving (3.13a) for $i = i_0$.

The complete scheme for generating the global estimate from the local ones is summarized as follows:

$$h_j(i+1) = F(i+1)h_j(i) + G_j(i+1)A_j(i)\hat{x}_j(i|i) \tag{3.15a}$$

$$\hat{x}(i|i) = \sum_{j=1}^{J}[P(i|i)M_j^T(i)P_j^{-1}(i|i)\hat{x}_j(i|i) + h_j(i)] \tag{3.15b}$$

The block diagram corresponding to these equations are shown in Figure 3.3. Notice that each of the local processors must send 2 vectors: $\hat{x}(i|i)$ and $h_j(i)$ to the central node.

The transmission requirement can be reduced if some of the computational burden of the central processor is transferred to the local ones. For example, if we define

$$\eta_j(i) = P(i|i)M_j^T(i)P_j^{-1}(i|i)\hat{x}_j(i|i) + h_j(i) \tag{3.16a}$$

then

$$\hat{x}(i|i) = \sum_{j=1}^{J}\eta_j(i) \tag{3.16b}$$

and the j-th local agent need to send only this vector to the mixing node - thus the transmission requirement has been reduced to 50%. Since all the variables in (3.16a) are generated at the node, $\eta_j(i)$ can be constructed locally at that node. The evolution of $\eta_j(i)$ can be established as follows:

$$\eta_j(i+1) = P(i+1|i+1)M_j^T(i+1)P_j^{-1}(i+1|i+1)\hat{x}_j(i+1|i+1) + h_j(i+1) \quad (3.17a)$$

Using the value of $h_j(i+1)$ from (3.13c) in this expression and after rearrangement gives

$$\begin{aligned}\eta_j(i+1) = {} & F(i+1)[h_j(i) + P(i|i)M_j^T(i)P_j^{-1}(i|i)\hat{x}_j(i|i)] \\ & + P(i+1|i+1)M_j^T(i+1)[P_j^{-1}(i+1|i+1)\hat{x}_j(i+1|i+1) \\ & - P_j^{-1}(i+1|i)\hat{x}_j(i+1|i)]\,. \end{aligned} \quad (3.17b)$$

But the expression in the second pair of brackets is

$$\begin{aligned} & P_j^{-1}(i+1|i+1)\hat{x}_j(i+1|i+1) - P_j^{-1}(i+1|i)\hat{x}_j(i+1|i) \\ & = [P_j^{-1}(i+1|i+1) - P_j^{-1}(i+1|i+1)K_j^d(i+1)H_j(i+1) - P_j^{-1}(i+1|i)]\hat{x}_j(i+1|i) \\ & \quad + P_j^{-1}(i+1|i+1)K_j^d(i+1)z_j(i+1) \\ & = K_j^c(i+1)z_j(i+1)\,; \end{aligned}$$

$$K_j^c(i+1) = P(i+1|i+1)C_j^T(i+1)R_j^{-1}(i+1)\,. \quad (3.17c)$$

Therefore (3.17a) becomes

$$\eta_j(i+1) = F(i+1)\eta_j(i) + K_j^c(i+1)z_j(i+1)\,. \quad (3.17d)$$

The initial condition for $\eta_j(i)$ is obtained from (3.16a):

$$\eta_j(i_0) = P(i_0|i_0)M_j^T(i_0)P_j^{-1}(i_0|i_0)\hat{x}_j(i_0|i_0) + h_j(i_0) \quad (3.17e)$$

The block diagram of this scheme is given in Figure 3.4. Notice that the explicit appearance of the local estimator has been totally eliminated. The structure of the central node is very simple - it simply adds the incoming vectors.

3.3 The Special Case of Identical Local and Global Models

In this section we consider the case examined by Speyer (1979) Specifically, we assume that the models used by the local processors are identical to the global model That is,

$$A = A_j,\ Q = Q_j,\ C_j = H_j,\ M_j = I\,. \tag{3.18}$$

In this case, the filtered estimate is generated as follows:

$$h_j(i+1) = F(i+1)h_j(i) + G_j(i+1)A(i)\hat{x}_j(i \mid i) \tag{3.19a}$$

$$\hat{x}(i \mid i) = \sum_{j=1}^{J}[P(i \mid i)P_j^{-1}(i \mid i)\hat{x}_j(i \mid i) + h_j(i)] \tag{3.19b}$$

where, as before,

$$F(i+1) = P(i+1 \mid i+1)P^{-1}(i+1 \mid i)A(i)$$
$$G(i+1) = P(i+1 \mid i+1)P^{-1}(i+1 \mid i)A(i)P(i \mid i)P_j^{-1}(i \mid i)A^{-1}(i)$$
$$\qquad - P(i+1 \mid i+1)P_j^{-1}(i+1 \mid i)\,.$$

The other scheme of generating $\hat{x}(i \mid i)$ using $\eta_j(i)$ remains unchanged in this special case and is identical to (3.16).

3.4 Summary and Conclusion

The decentralized estimation scheme for the discrete time case has been presented in this chapter whereas that for the continuous time case was presented in Chapter 2. In Section 3 2 the problem has been formulated in a more general setting: here we have assumed that the local processor models are not necessarily the same as the true global model except that the sensor noise statistics is same for both the models. The globally optimum filtered states of the target can be constructed as follows. raw observed data from

various sensor locations can be transmitted to a central (mixing) node where all the data can be processed for the global estimate. Alternatively, raw observed data can be processed locally and then the local estimates are sent to the mixing node where these are combined into the global estimate. Throughout this report we have emphasized that transmitting processed data from the sensor locations to a mixing node is superior to transmitting raw data. Along this line, we have shown in this section that the globally optimal estimate can be constructed at a central node from two processed data vector from each of the local processors, these are locally optimum estimate vector and another data dependent vector that evolves recursively according to Equation (3.14a). This scheme is shown in Figure 3 3. This scheme can be improved further. the transmission requirement can be reduced by 50% by transferring some of the computational burden of the central processor to the local ones. In this improved scheme, each local processor need to send only one data vector to the central node, this data vector is generated recursively at the local node according to Equation (3.17d). This scheme is shown in Figure 3.4. The case of identical global and local models is dealt with in Section 3.3 - it is observed that much simplifications do not occur in this case.

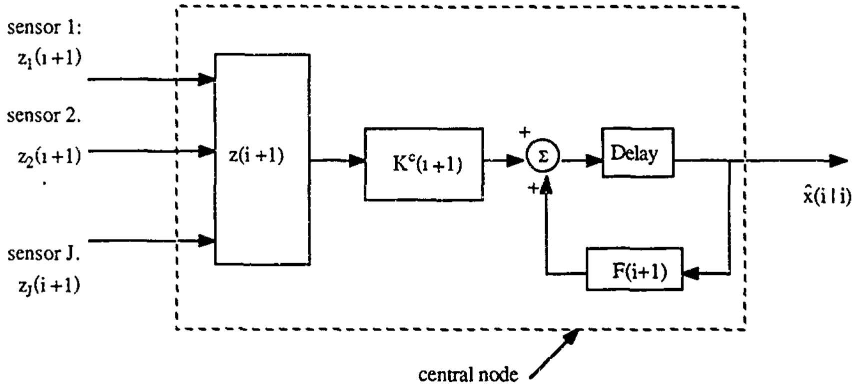


Figure 3.1: Centralized Estimation

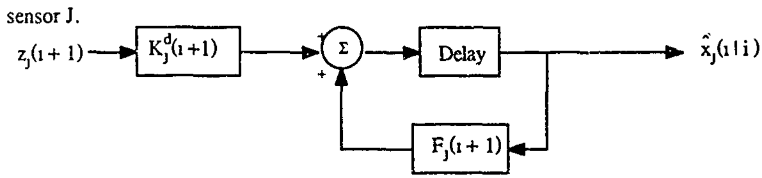


Figure 3.2: Local Estimator: j-th Sensor

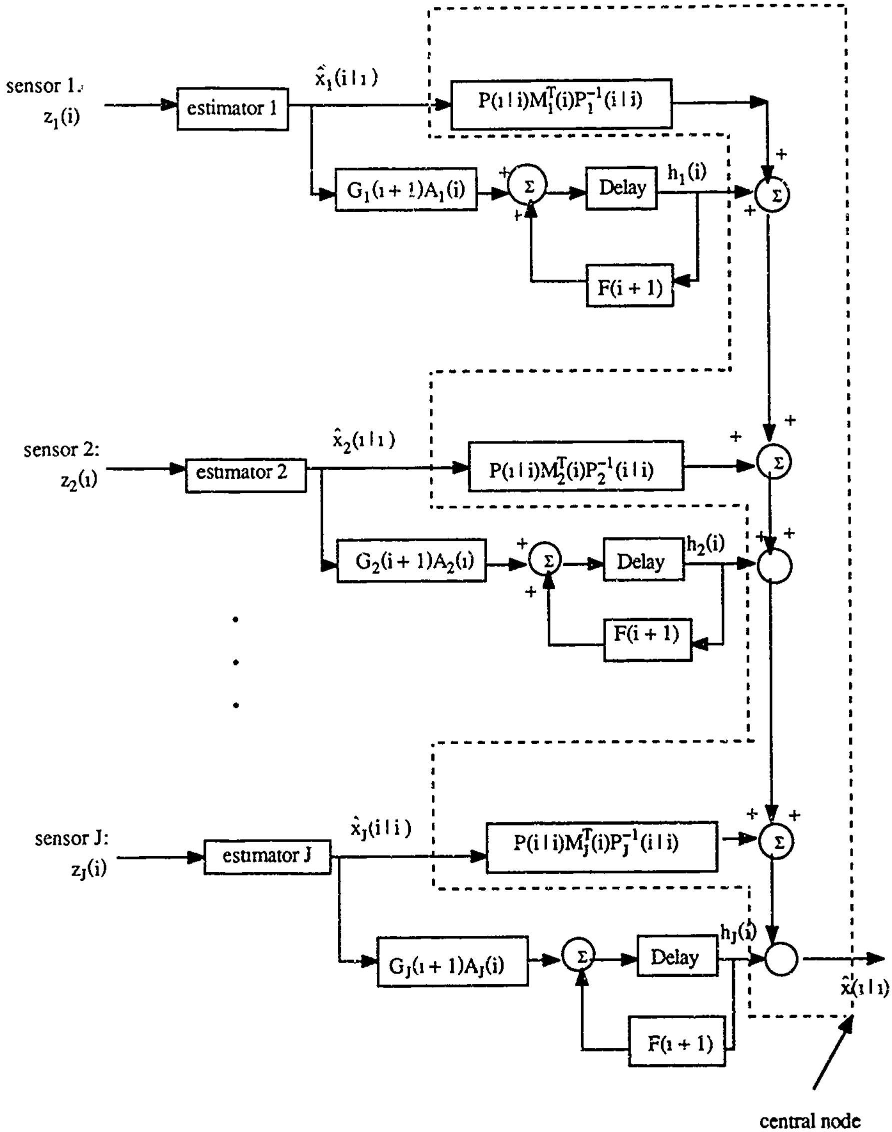


Figure 3.3: Decentralized Estimation: A Combination of Local and Central Processing

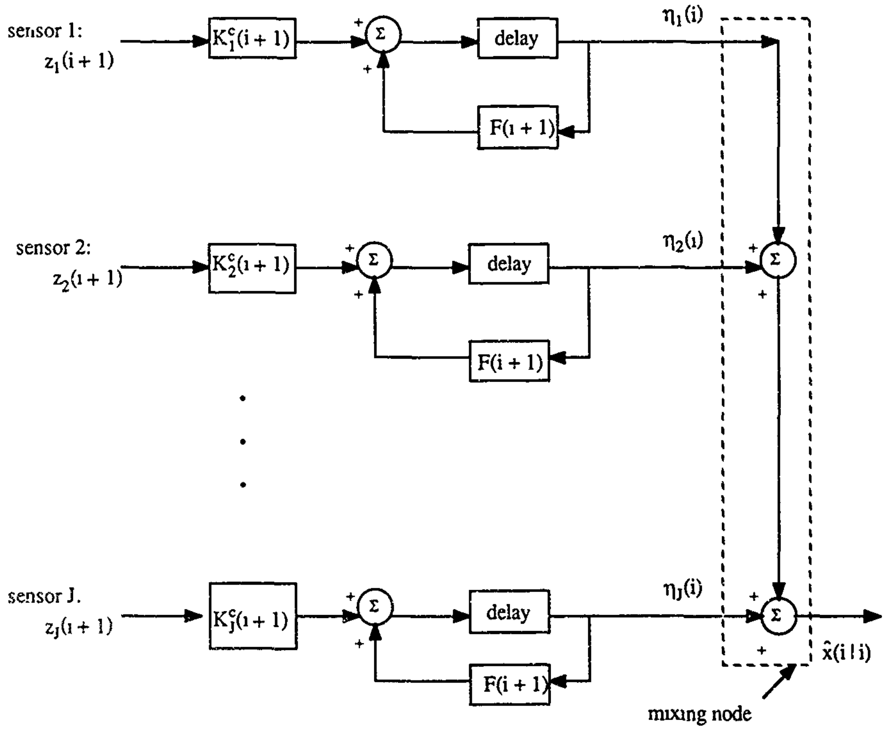


Figure 3.4: Decentralized Estimation: A Simpler Local Processor Structure and Reduced Transmission

CHAPTER 4

DECENTRALIZED ESTIMATION IN PRESENCE OF CORRELATED SENSOR NOISE

4.1 Introduction

In Chapters 2 and 3, the decentralized estimation problem was solved for the case of uncorrelated noise. In this chapter, a simpler version of the same problem will be solved for the case of correlated sensor noise. Because of the importance of this problem, we have tried to make this a self contained chapter.

The same scenario as that of the previous chapters will be adopted here: multiple sensors that are possibly located at dispersed geographical locations are observing a single SDI target. Attached with each sensor is a data processing capability. A sensor along with its own data processing capability is also known as a local agent or a local processor. The true state of the target is also known as the global state and the ultimate goal is to obtain the optimal estimate of this state by utilizing all the sensor measurements. The dynamic model of the target state may or may not be available to the local agents. We have discussed in the previous chapters that the global estimation problem can be solved in two alternative ways. First, all the sensors can transmit their raw observations to a central node (also known as a fusion center or mixing node) to form an augmented observation vector and a centralized Kalman filter can then extract the optimal state from this vector. Alternatively, each local agent, by utilizing its own model about the global state, can process its own observations on site to form the local optimal estimate conditional on its measurement history and then transmit this estimate to the fusion center. In the fusion center, the incoming estimates from various nodes are combined in an appropriate fashion to construct the globally optimal estimate - this is known as a decentralized or distributed estimation scheme. It has been

mentioned earlier that the latter alternative is superior from the system survivality considerations In the centralized scheme, the system performance is lost completely in the event of the central mode failure On the other hand, in the decentralized scheme, a graceful degradation of the system performance takes place in the event of a node failure In order to build an ideal redundancy into the system, each of the local nodes should be designed as a fusion center in the sense that each node will receive estimates from all other nodes and generate an identical globally optimal estimate of the state. This can be mechanized if all the sensors are communicating in a broadcast mode. As in the previous chapters, we assume that there is no appreciable transmission or computational delay. Therefore, after an observation is made in the discrete time case, the nodes will compute the locally optimal estimate and transmit it before the next observation is made. Furthermore, we also assume that all the sensors are engaged in time-synchronized measurements For the continuous time case, we are assuming that computation at and transmission from each node take place instantaneously. We will consider that the SDI target is governed by linear time varying dynamics and measurement equations.

Recently, there has been a great deal of interest in the decentralized estimation. Speyer (1979) first formulated the problem in an LQG framework - he solved the decentralized estimation as a prelude to a decentralized control problem He assumed that the global and local models were identical. Willsky et al. (1982) solved the problem for a more general case of nonidentical global and local models - they have also solved the distributed smoothing problem in this paper. Recently, Hashemipour, Roy and Lamb (1984) have expanded the distributed estimation problem to include collocated and non-collocated sensors, and have provided implementation details for the discrete time case. In all of these works, the authors have assumed that the measurement noise processes are uncorrelated from sensor to sensor. The primary purpose of this chapter is to develop a distributed estimation algorithm when the sensor noises are correlated. The underlying

assumption is that the local agents have the full knowledge of the global model and various statistics of the system.

The organization of this chapter is as follows. In Section 4.2, the important results about the decentralized estimation for the case of uncorrelated sensor noise are reviewed. The case of correlated sensor noise in the continuous time domain is presented in Section 4.3 Several schemes for optimal estimation are presented in this section: some of these schemes employ centralized processing algorithms. The last of these schemes is truly a distributed one. We will consider the continuous time case only, the corresponding results for the discrete time case will be analyzed in Phase II. Finally, a summary and conclusion is provided in Section 4 4.

4.2 A Review of Distributed Estimation Problem in Presence of Uncorrelated Sensor Noise: Continuous Time Case

In this section we assume that the measurement noises are uncorrelated across the sensors. Suppose that there are J sensors tracking a single SDI target. The dynamic model of the target and the measurement equations are assumed as follows:

$$x(t) = A(t)x(t) + w(t) \ ; \qquad t \geq t_0 \tag{4.1a}$$

$$z_j(t) = c_j(t)x(t) + v_j(t) \ ; \qquad j = 1,2,...J. \tag{4.1b}$$

Various statistics relating to this model are:

$$E\,[x(t_0)] = \bar{x}(t_0), \quad E\,[(x(t_0) - \bar{x}(t_0))\,(x(t_0) - \bar{x}(t_0))^T] = \Sigma(t_0)\,,$$

$$E\,[w(t)] = 0, \quad E\,[w(t)w^T(\tau)] = Q(t)\delta(t-\tau)$$

$$E\,[v_j(t)] = 0, \quad E[v_j(t)v_k^T(\tau)] = R_j(t)\delta(t-\tau)\delta_{jk}, \quad R_j(t) = R_j^T(t) > 0$$

We also assume that $w(t)$ is independent of $x(t_0)$ and of $v_j(t)$ for all $j = 1, 2, \ldots J$. It is also assumed that $v_j(t)$ is independent of $x(t_0)$.

Since the underlying assumpuon is that the local processors have the full knowledge of the global model, the state model at the j-th sensor can be written as

$$\dot{x}_j(t) = A(t)x_j(t) + w(t) , \tag{4.2a}$$

with

$$E[x_j(t_0)] = \bar{x}_j(t_0), \quad E[(x_j(t_0) - \bar{x}_j(t_0)) (x_j(t_0) - \bar{x}_j(t_0))^T] = \Sigma_j(t_0) .$$

The measurement equation and other related statistics are same as above.

The globally optimal filtered state conditional on the measurement history of all the sensors is given by

$$\hat{x}(t) = E [x(t) \mid z_1(\zeta), z_2(\zeta) \ldots z_J(\zeta) ; \quad t_0 \leq \zeta \leq t] .$$

This estimate is propagated as

$$\dot{\hat{x}}(t) = A(t)\hat{x}(t) + \sum_{j=1}^{J} K_j^{cu}(t) [z_j(t) - C_j(t)\hat{x}(t)] , \tag{4.3a}$$

where

$$K_j^{cu}(t) = P(t)C_j^T(t)R_j^{-1}(t) , \tag{4.3b}$$

and the estimate error covariance matrix $P(t)$ is the solution of the Riccati equation

$$P(t) = A(t)P(t) + P(t)A^T(t) + Q(t) - \sum_{j=1}^{J} P(t)C_j^T(t)R_j^{-1}(t)C_j(t)P(t) \tag{4.3c}$$

$$P(t_0) = \Sigma(t_0) .$$

Clearly the mechanization of (4.3) requires that all the sensor measurements be transmitted to a central node where $\hat{x}(t)$ will be generated We have used the notation $K_j^{cu}(t)$ to imply that this is the centralized Kalman gain when the sensor noises are uncorrelated.

If the nodes are allowed to process their own observations locally, the local optimal estimate evolves as.

$$\dot{\hat{x}}_j(t) = A(t)\hat{x}_j(t) + K_j^{du}(t)\,[z_j(t) - C_j(t)\hat{x}_j(t)] \tag{4.4a}$$

where

$$K_j^{du}(t) = P_j(t)C_j^T(t)R_j^{-1}(t) \tag{4.4b}$$

and the error covariance matrix $P_j(t)$ can be computed from the Riccati equation

$$\dot{P}_j(t) = A(t)P_j(t) + P_j(t)A^T(t) + Q(t) - P_j(t)C_j^T(t)R_j^{-1}(t)C_j(t)P_j(t) \tag{4.4b}$$

$$P_j(t_0) = \Sigma_j(t_0) .$$

Here, $K_j^{du}(t)$ stands for the decentralized or localized Kalman gain associated with the j-th node in presence of uncorrelated sensor noise. The problem of decentralized or distributed estimation is to reconstruct $\hat{x}(t)$ in terms of $\hat{x}_j(t)$, j = 1, 2, ...J of (4.4a) Willsky et al. (1982) has shown that $\hat{x}(t)$ can be constructed at a central node dynamically from the local estimates as follows:

$$\dot{\zeta}(t) = F(t)\zeta(t) + \sum_{j=1}^{J} K_j(t)\hat{x}_j(t) \tag{4.5a}$$

$$\hat{x}(t) = \zeta(t) + \sum_{j=1}^{J} P(t)P_j^{-1}(t)\hat{x}_j(t) \tag{4.5b}$$

where

$$F(t) = A(t) - \sum_{j=1}^{J} P(t)C_j^T(t)R_j^{-1}(t)C_j(t) , \tag{4.5c}$$

$$K_j(t) = [P(t)P_j^{-1}(t) - I]Q(t)P_j^{-1}(t) . \tag{4.5d}$$

These equations show that each node j must send to the central node its local estimate $\hat{x}_j(t)$ and the associated covariance matrix $P_j(t)$. The structure of the central node can be simplified using Speyer's (1979) form where some of the computational burden of the central node is transferred to the local nodes. At each node j, a data dependent vector $h_j(t)$ is generated as follows:

$$\dot{h}_j(t) = F(t)h_j(t) + K_j(t)\hat{x}_j(t) . \tag{4.6a}$$

Then, at the central node, the global estimate is constructed as

$$\hat{x}(t) = \varphi_F(t, t_0)\hat{x}(t_0) + \sum_{j=1}^{J} [P(t)P_j^{-1}(t)\hat{x}_j(t) + h_j(t)] ,$$

where $\phi_F(t, t_0)$ is the transition matrix associated with the dynamics matrix F(t) The complexity of the central node has been reduced at the expense of higher transmission requirements - now each node j must send to the central node 2 vectors the locally optimal estimate $\hat{x}_j(t)$ and a data dependent vector $h_j(t)$. It has been shown in Chapter 2 that the transmission requirements can further be reduced to only one data dependent vector.

The key to the distributed estimation algorithm (4.5) and (4.6) is the fact that the global estimate $\hat{x}(t)$ can be decomposed in terms of the locally computed data dependent vectors as shown in these equations This decomposition has been possible because the sensor noises are uncorrelated It 's the purpose of the next section to show how this decomposition can be achieved when the sensor noises are correlated across the sensors

4.3 Distributed Estimation with Correlated Sensor Noise: Continuous Time Case

Consider again the dynamic model (4.1) having the same process noise and initial condition. But this time we assume that the sensor noises are correlated Specifically we assume

$$E\,[v_j(t)] = 0, \quad E\,[v_j(t)v_k^T(\tau)] = R_{jk}(t)\delta(t-\tau) \tag{4.7}$$

As before, we assume that the process noise $w(t)$ is independent of $x(t_0)$ and of $v_j(t)$ for all $j = 1, 2, \ldots J$. Also, $v_j(t)$ is independent of $x(t_0)$. The optimal estimation problem is first solved for the centralized case, i.e., where the raw measurements from all the sensors are transmitted to a central node.

A Centralized Solution for the Optimal Estimation Problem

The optimal filtered estimate is given by

$$\hat{x}(t) = E\,\{x(t) \mid z_j(\zeta)\,; \quad t_0 \le \zeta \le t\,,\, j = 1, 2, \ldots J\,\}$$

and can be computed as follows in a straightforward way Define

$$z(t) \triangleq \begin{bmatrix} z_1(t) \\ z_2(t) \\ \cdot \\ \cdot \\ z_J(t) \end{bmatrix} ; \quad C(t) \triangleq \begin{bmatrix} C_1(t) \\ C_2(t) \\ \cdot \\ \cdot \\ \cdot \\ C_J(t) \end{bmatrix} ; \quad v(t) = \begin{bmatrix} v_1(t) \\ v_2(t) \\ \cdot \\ \cdot \\ v_J(t) \end{bmatrix} . \tag{4.8a}$$

Then all the sensor measurements can be compactly represented as

$$z(t) = C(t)x(t) + v(t) \tag{4.8b}$$

Here v(t) is a white noise process whose covariance matrix is given by

$$E\,[v(t)v^T(\tau)] = R(t)\delta(t-\tau)\,; \quad R(t) = R^T(t) > 0 \tag{4.9a}$$

where

$$R(t) = \text{block}\,[R_{ij}(t)\,; \quad i, j = 1, 2, ..J] \tag{4.9b}$$

Clearly, for the case of uncorrelated sensor noise, R(t) is a block diagonal matrix and it is this block diagonal structure of the R(t) matrix that has led to the decomposition of the global estimates as in Equations (4.5) - (4.6) For the correlated sensor noise this is a "full" matrix. However, irrespective of the structure of this matrix, the optimal filtered estimate is given by

$$\hat{x}(t) = E[x(t) \mid z(\zeta)\,; \quad t_0 \le \zeta \le t]$$

which can be computed from the standard Kalman filter.

$$\dot{\hat{x}}(t) = A(t)\hat{x}(t) + K^{cc}(t)[z(t) - C(t)\hat{x}(t)]\,, \tag{4.10a}$$

$$K^{cc}(t) = P(t)C^T(t)R^{-1}(t)\,, \tag{4.10b}$$

and the covariance matrix P(t) obeys

$$P(t) = A(t)P(t) + P(t)A^T(t) + Q(t) - P(t)C^T(t)R^{-1}(t)C(t)P(t) \tag{4.10c}$$

$$P(t_0) = \Sigma(t_0) .$$

$C(t)$ and $R(t)$ have been defined in (4.8a) and (4.9a) respectively. Here $K^{cc}(t)$ stands for the centralized Kalman gain for the correlated sensor noise The scheme of (4 10) can be mechanized at a central node where all the sensor measurements $z_j(t)$ are transmitted to form the aggregate measurement vector $z(t)$

Alternatively, this vector $z(t)$ can be transformed into pseudo-measurement vector $\tilde{z}(t)$ such that the measurement noise associated with this new vector corresponding to various sensors is uncorrelated To obtain $\tilde{z}(t)$, decompose $R(t)$ in such a way that

$$R(t) = R^{1/2}(t)R^{T/2}(t) \tag{4.11a}$$

where $R^{1/2}(t)$ is a positive definite square root of $R(t)$. Then

$$\begin{aligned} \tilde{z}(t) &= R^{-1/2}(t)z(t) \\ &= \tilde{C}(t)x(t) + \tilde{v}(t) \end{aligned} \tag{4.11b}$$

where

$$\tilde{C}(t) = R^{-1/2}(t)C(t)\,, \quad \tilde{v}(t) = R^{-1/2}v(t)\,. \tag{4.11c}$$

Clearly, the covariance matrix $\tilde{R}(t)$ of $\tilde{v}(t)$ is the identity matrix for all t and satisfies therefore the new measurement noise $\tilde{v}(t)$ is uncorrelated across the sensors. It can be verified in a straightforward way that the Riccati equations associated with the Kalman filters for the original form (4.10) and the new measurement coordinate systems (4.11) are the same Therefore the optimal estimate in the new measurement system is given by

$$\dot{\hat{x}}(t) = A(t)\hat{x}(t) + \sum_{j=1}^{J} \tilde{K}_j^{cc}(t)\,[\tilde{z}_j(t) - \tilde{C}_j(t)\hat{x}(t)]\,, \tag{4.12a}$$

$$\tilde{K}_j^{cc}(t) = P(t)\tilde{C}_j(t) \tag{4.12b}$$

where P(t) is obtained from (4.10c) and $\tilde{C}_j(t)$ is the j-th submatrix of C(t) corresponding to the j-th sensor. Similarly, $\tilde{z}_j(t)$ is the corresponding j-th subvector of $\tilde{z}(t)$. This scheme looks the same as for the uncorrelated sensor noise as shown in (4 3), but yet can not be implemented in a decentralized framework. The reasons are as follows. although $\tilde{C}_j(t)$ and $\tilde{K}_j^{cc}(t)$ can be computed at the j-th node, but not $\tilde{z}_j(t)$. In fact $\tilde{C}_j(t)$ and $\tilde{K}_j^{cc}(t)$ can be computed and stored apriori. In order to compute $\tilde{z}_j(t)$, z(t) must be formed first at that node and therefore must receive the raw observations from all other nodes. Once z(t) is formed, it is to be multiplied by $R^{-1/2}(t)$ and $\tilde{z}_j(t)$ is the j-th subvector of the resulting vector. Our goal is to process the raw observation data locally and then transmit the processed data to a central node.

Distributed Optimal Estimation

In this section we will show how each node will generate a data dependent vector from its own observations only and without utilizing data from other sensors. These vectors from all the nodes will be communicated to a mixing node where these will be fused into the globally optimal estimate. The main result is presented in the following fact.

Fact 1

Suppose at each node j, a data dependent vector $q_j(t)$ is generated as follows:

$$\dot{q}_j(t) = F(t)q_j(t) + K_j^{cc}(t)z_j(t)\,; \quad q_j(t_0) = 0, \tag{4.13a}$$

where

$$F(t) = A(t) - K^{cc}(t)C(t) \tag{4.13b}$$

and $K_j^{cc}(t)$ is the j-th submatrix from a compatible decomposition of $K^{cc}(t)$ of the form

$$K^{cc}(t) = [K_1^{cc}(t), K_2^{cc}(t) \ ..K_J^{cc}(t)] \tag{4.13c}$$

Then

$$\hat{x}(t) = \phi_F(t, t_0)\hat{x}(t_0) + \sum_{j=1}^{J} q_j(t) \,. \tag{4.13d}$$

Here $K^{cc}(t)$ is the centralized Kalman gain given in (4.10b) and C(t) is the measurement matrix associated with the aggregate measurement vector z(t) as shown in (4.8); $K_j^{cc}(t)$ is the appropriate block of $K^{cc}(t)$ corresponding to the j-th sensor and $\phi_F(t, t_0)$ is the transition matrix associated with F(t).

This is a simple yet remarkable result. Since the underlying assumption is that the local agents have the full knowledge of the global model, all the variables of (4.13b) are available to the j-th processor. $K^{cc}(t)$ can be locally computed or can be transmitted from a central coordinator - in either case it can be stored apriori at the j-th sensor The same comment applies to the construction of F(t). Moreover, each local processor can be initialized to zero thus avoiding the requirement of a complicated initial setting. Therefore $q_j(t)$ is truly a node dependent vector and can be constructed from the observations at that node only.

An implementation of this scheme is shown in Figure 4.1. It can be seen from this figure that, the node j does not have an impiicit structure of a Kalman filter, neither the central node has a Kalman filter in it. Each node sends only one data dependent vector $q_j(t)$

to the central node - thus the transmission requirement is minimal. The structure of the central node is also quite simple; if $\hat{x}(t_0) = 0$, it is merely an adder.

Proof

The proof is straightforward. We will show that $\hat{x}(t)$ generated this way indeed satisfies the standard Kalman filter equations given in (4.10). Clearly, since $q_j(t_0) = 0$,

$$q_j(t) = \int_{t_0}^{t} \phi_F(t, \tau) K_j^{cc}(\tau) z_j(\tau) d\tau \tag{4.14a}$$

Substituting $q_j(t)$ in (4.13d), we get

$$\hat{x}(t) = \phi_F(t, t_0)\hat{x}(t_0) + \sum_{j=1}^{J} \int_{t_0}^{t} \phi_F(t, \tau) K_j^{cc}(\tau) z_j(\tau) d\tau . \tag{4.14b}$$

Differentiating both sides,

$$\dot{\hat{x}}(t) = F(t)\hat{x}(t) + \sum_{j=1}^{J} K_j^{cc}(t) z_j(t) . \tag{4.14c}$$

Finally, rearrangement of this equation gives

$$\dot{\hat{x}}(t) = A(t)\hat{x}(t) + K^{cc}(t)\,[z(t) - C(t)\hat{x}(t)] \tag{4.15a}$$

which is the same as (4.10). ■

There are several interpretations of $q_j(t)$. For the case of uncorrelated sensor noise, it is the locally optimal estimate. When the sensor noises are correlated, $q_j(t)$ can be regarded as a sub-optimal estimate of the global state as viewed by the local agent. This vector will be studied in detail in Phase II.

4.4 Conclusions

The decentralized estimation problem for the correlated sensor noise has been solved in this chapter. The underlying idea of the decentralized estimation is that the observations are processed locally by each node and then the processed data is sent to a central node where the globally optimal estimate is constructed The problem with uncorrelated sensor noise has been reviewed first in Section 4.2, because most of these results were used to derive corresponding results for the correlated sensor noise case. It has been shown in Section 4.3 that for the later case, each node must generate a data dependent vector $q_j(t)$ which is then transmitted to the central node and fused with similar vectors from other nodes into the globally optimal estimate. The detailed properties of this vector are yet to be investigated.

We have solved this problem for a special case of identical global and local models When the models differ, the analysis is very difficult and will be taken up in the forthcoming Phase II.

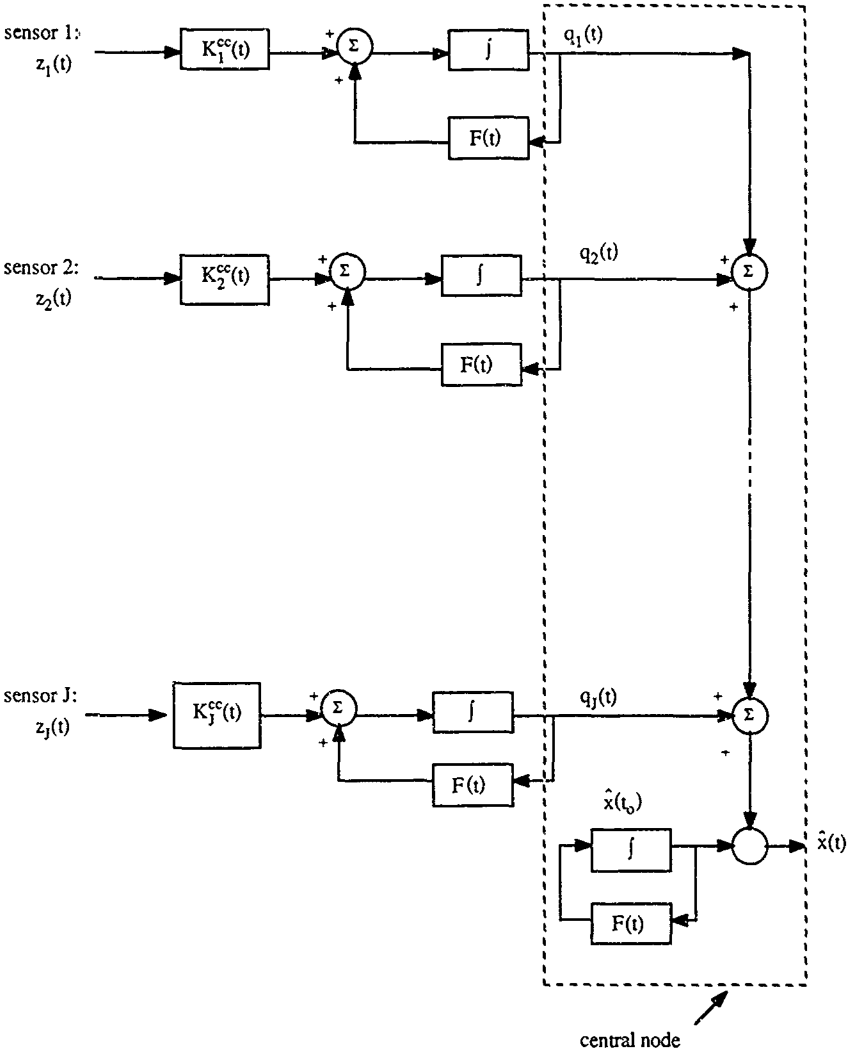


Figure 4.1: Distributed Estimation with Correlated Sensor Noise

CHAPTER 5

DEVELOPMENT AND APPLICATION OF DECENTRALIZED MODIFIED GAIN EXTENDED KALMAN FILTER TO SDI TRACKING

5.1 Introduction

In this chapter, the results of decentralized estimation of the previous chapters are extended to a special class of nonlinear measurements encountered in the SDI tracking problem The IR sensors used in the SDI system typically provide information on bearings of the targets Radar and Laser trackers can provide additional information on range and range-rate If the motion of the target is described in a rectangular coordinate system, the above measurements are related nonlinearly to the state variables of the target

The approach used to extend the results of the linear decentralized Kalman Filter to the above case is to use Modified Gain Extended Kalman Filters.

In this chapter, a new nonlinear filter, whose structure is similar to that of the Extended Kalman Filter (EKF), is applied to the estimation problem using a strapdown seeker. This nonlinear filter is based upon the system nonlinearities being members of a special class of functions called modifiable. The essential idea is that the difference between the nonlinear function at an unknown state and a known state is equal to a linear function in the difference between the unknown and known state operated upon by a matrix function composed only of the known state and the measurement functions. Thus, this special class possesses a natural linearity. Although this class is quite small, it does include several of the nonlinearities present in the SDI tracking problem. The theoretical analysis of centralized estimators containing this type of nonlinearity with interesting applications are given in Song and Speyer (1984), Safonov and Athans (1978) The theoretical results show that if no noise inputs are present, the filter, acting as an observer, is globally stable

If noise is present, the centralized filter is shown with additional conditions to be stochastically exponentially bounded The convergence of the estimates to unbiased values has been indicated mostly by simulation (Song and Speyer, 1984) but for the parameter identification problem can be shown analytically (Safonov and Athans, 1978)

Since the class of nonlinear functions that are modifiable may not include all the nonlinearities present in a state estimation problem, a filter structure involving both modifiable and nonmodifiable functions is possible for sophisticated applications such as SDI Consequently, those nonlinearities which are modifiable are included in the estimation algorithm in their modifiable form and those that are not modifiable are included as they would be in a standard EKF. In the SDI tracking problem, the essential or most important nonlinearities belong to the class of modifiable nonlinearities. The essential, nonlinearities are the functional forms of the elevation and azimuth angles, the range, and the range rate formulated in rectangular coordinates The misalignment errors produce nonlinearities which are included by the linearization method of the standard EKF. The resulting filter is called the modified gain EKF (MGEKF), since only the gain calculations are somewhat different from those of the EKF but not the estimator structure

If both angles and range and range rate are measured, then the anticipated advantages of the MGEKF over the EKF may not be clearly seen since in both schemes the large number of measurements allows good estimation of the states However, if only angle information is present which is the case with IR measurements, then it is anticipated that significant improvement in filter response and stability should be evident in the MGEKF formulation

5.2 Modifiable Functions in the SDI Tracking Problem

The measurements considered here are the elevation and azimuth angles, range and range rate or any subset of these measurements. These measurements are put in modifiable form first. Then the modification to the gain calculation is given.

Modifiable Form of Measurements

Let

$$X \stackrel{\Delta}{=} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{5.1}$$

denote the SDI target position in inertial coordinate frame and

$$\underset{-B}{X} \stackrel{\Delta}{=} \begin{bmatrix} X_B \\ Y_B \\ Z_B \end{bmatrix} = T \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{5.2}$$

denote the target position in a body axis system centered at the sensored platform where T denotes the transformation matrix between the two coordinate frames.

The measurements are given in deterministic form as <u>angle measurement functions</u>:

$$\begin{bmatrix} \varepsilon_Y \\ \varepsilon_Z \end{bmatrix} = \begin{bmatrix} K_Y & K_{YZ} \\ K_{YZ} & K_Z \end{bmatrix} \begin{bmatrix} \tan^{-1}(Y_B/X_B) \\ \tan^{-1}\left(\dfrac{-Z_B}{\sqrt{X_B^2+Y_B^2}}\right) \end{bmatrix} + \begin{bmatrix} G_Y \\ G_Z \end{bmatrix} \tag{5.3}$$

where $G^T = [G_Y, G_Z]^T$ is the glint noise state vector

$$\text{range function: } R = \sqrt{X^2 + Y^2 + Z^2} = \sqrt{X_B^2 + Y_B^2 + Z_B^2}\ , \tag{5.4}$$

$$\text{range rate function: } \dot{R} = \frac{XV_X + YV_Y + ZV_Z}{R} \tag{5.5}$$

where $\underline{V} \overset{\Delta}{=} [V_x, V_y, V_z]^T$ is inertial velocity vector and $\underline{K} \overset{\Delta}{=} [K_Y, K_{YZ}, K_T]^Z$ denotes a vector sensor scale factor parameters

The measurement function given by (5.3) to (5.5) are represented as the vector function

$$z^* \overset{\Delta}{=} h(\underline{X}, \underline{G}, \underline{V}, \underline{K}) \tag{5.6}$$

The actual measurement is given by $z = z^* + v$ where v is white zero mean Gaussian noise The global exponential boundedness for the stochastic case are obtained by assuming that certain finite gain operators associated with the additive measurement noise do not destabilize the system Results of this sort are given in Song and Speyer (1985) with respect to a constant gain EKF. The results of this chapter extend these results to time varying gain.

5.3 The Definition of a Modifiable Function

The measurement function is modifiable if

$$h(\underline{X}, \underline{G}, \underline{V}, \underline{K}) - h(\overline{\underline{X}}, \overline{\underline{V}}, \overline{\underline{G}}, \overline{\underline{K}}) = F(z^*, \underline{X}, \underline{G}, \underline{V}, \underline{K}) \begin{bmatrix} \underline{X} - \overline{\underline{X}} \\ \underline{V} - \overline{\underline{V}} \\ \underline{G} - \overline{\underline{G}} \\ \underline{K} - \overline{\underline{K}} \end{bmatrix} \tag{5.7}$$

where $\underline{X}, \underline{G}, \underline{V}, \underline{K}$ are unknown values but $\overline{\underline{X}}, \overline{\underline{V}}, \overline{\underline{G}}, \overline{\underline{K}}$ are estimated values of the state and are thereby known. Note that F will be evaluated using z rather than z* in implementing the filter. Furthermore, note that for continuous functions

$$F(z^*, \overline{\underline{X}}, \overline{\underline{V}}, \overline{\underline{G}}) \Rightarrow \frac{\partial h}{\partial(\underline{X}, \underline{G}, \underline{V}, \underline{K})}\Big|_{(\underline{X}, \underline{G}, \underline{V}, \underline{K}) = (\overline{\underline{X}}, \overline{\underline{V}}, \overline{\underline{G}}, \overline{\underline{K}})} \qquad (5.8)$$

Therefore, in the limit the MGEKF converges to the EKF, as $(\underline{X}, \underline{G}, \underline{V}) \Rightarrow (\overline{\underline{X}}, \overline{\underline{V}}, \overline{\underline{G}})$ This should be used as a check on the numerical implementation

5.4 List of Preliminary Functions

The following functions are required in implementing the MGEKF.

$$\overline{az} = \tan^{-1}\frac{\overline{Y}_B}{\overline{X}_B} \qquad (5\ 9)$$

$$\overline{el} = \tan^{-1}\frac{-\overline{Z}_B}{\sqrt{\overline{X}_B^2 + \overline{Y}_B^2}} \qquad (5.10)$$

$$H(az, el) = \begin{bmatrix} \sin az, & -\cos az, & 0 \\ \sin el \cos az, & \sin el \sin az, & \cos el \end{bmatrix} \qquad (5.11)$$

$$E = \begin{bmatrix} \dfrac{\tan^{-1}\alpha}{(\cos az\, \overline{X}_B + \sin az\, \overline{Y}_B} & 0 \\ 0 & \dfrac{\tan^{-1}\beta}{[\cos el\, (\overline{X}_B^2 + \overline{Y}_B^2)^{1/2} - \sin el\, \overline{Z}_B]\, \beta} \end{bmatrix} \qquad (5.12)$$

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \dfrac{1}{\cos az\, \overline{X}_B + \sin az\, \overline{Y}_B} & 0 \\ 0 & \dfrac{1}{\cos el (\overline{X}_B^2 + \overline{Y}_B^2)^{1/2} - \sin el\, \overline{Z}_B} \end{bmatrix} H\underline{\overline{X}}_B \qquad (5.13)$$

$$\begin{bmatrix} az \\ el \end{bmatrix} = \overline{K}^{-1}\begin{bmatrix} \varepsilon_Y \\ \varepsilon_Z \end{bmatrix} - \overline{K}^{-1}\overline{G} \qquad (5.14)\ *$$

$$\begin{bmatrix} az \\ el \end{bmatrix} = \overline{K}^{-1} \begin{bmatrix} \varepsilon_Y \\ \varepsilon_Z \end{bmatrix} - \overline{K}^{-1}\overline{G} \qquad (5.14)^{*}$$

$$a_1 = \frac{\cos az - \cos \overline{az}}{az - \overline{az}} \qquad (5\ 15)\ \dagger$$

$$a_2 = \frac{\sin az - \sin \overline{az}}{az - \overline{az}} \qquad (5.16)\ \dagger$$

$$e_1 = \frac{\cos el - \cos \overline{el}}{el - \overline{el}} \qquad (5.17)\ \dagger$$

$$e_2 = \frac{\sin el - \sin \overline{el}}{el - \overline{el}} \qquad (5.18)\ \dagger$$

$$\cos az = \overline{X}_B \Big/ \sqrt{\overline{X}_B^2 + \overline{Y}_B^2} \qquad (5.19)$$

$$\sin az = \overline{Y}_B \Big/ \sqrt{\overline{X}_B^2 + \overline{Y}_B^2} \qquad (5.20)$$

$$\cos \overline{el} = \frac{\overline{Z}_B}{\overline{R}} \qquad (5.21)$$

$$\sin \overline{el} = \frac{\sqrt{\overline{X}_B^2 + \overline{Y}_B^2}}{\overline{R}} \qquad (5.22)$$

$$\overline{R} = \sqrt{\overline{X}_B^2 + \overline{Y}_B^2 + \overline{Z}_B^2} \qquad (5.23)$$

$$F_A = -\overline{K}\, EHT^{-1} \qquad (5.24)$$

$$F_M = \begin{vmatrix} az, el, 0 \\ 0, az, el \end{vmatrix} \qquad (5.25)$$

---

* Note ε_Y, ε_Z, and R are replaced by their associated noisy measurements when the MGEKF is implemented

† If (5.15) to 5 20) have numerical difficulty due to division by a small number then approximate as $a_1 = \sin\theta$, $a_2 = \cos\theta$, $e_1 = \sin\phi$, $e_2 = \cos\phi$ where $\theta = (1/2)(az + \overline{az})$ and $\phi = (1/2)(el + \overline{el})$

$$F_R = \left\{ \begin{array}{l} (\cos az \cos el, \sin az \cos el, -\sin el) - (\overline{X}_B \cos el\, a_1 \\ + \overline{Y}_B \cos el\, a_2, \overline{X}_B \cos az\, e_1 + \overline{Y} \sin \overline{az}\, e_2 - \overline{Z}\, e_2)\, EH \end{array} \right\} T^{-1} \qquad (5.26)$$

$$F_{RV} = (\cos el \cos az, \cos el \sin az, -\sin el)\, T^{-1} \qquad (5.27)$$

$$F_{RR} = \left\{ \left(\frac{\overline{V}_X}{R}, \frac{\overline{V}_Y}{R}, \frac{\overline{V}_Z}{R} \right) - \left(\frac{\overline{XV}_X + \overline{YV}_Y + \overline{ZV}_Z}{R\,\overline{R}} \right) F_R \right\} \qquad (5.28)$$

5.5 The Modifiable Forms for the Measurements

The measurement function given by Eq. (5.6) is only approximately modifiable. The approximate modifiable form given by Eq (5.7) is presented below The angular measurements become approximately modifiable as

$$\begin{bmatrix} \varepsilon_y \\ \varepsilon_y \end{bmatrix} - \begin{bmatrix} \varepsilon_y \\ \varepsilon_y \end{bmatrix} = F_A(\underline{x}-\overline{\underline{x}}) + \underline{G} - \overline{\underline{G}} + F_M(\underline{K}-\overline{\underline{K}}) \qquad (5.29)$$

where

$$\underline{K} = \begin{bmatrix} K_Y \\ K_{YZ} \\ K_Z \end{bmatrix} \qquad (5\ 30)$$

The range measurement becomes approximately modifiable as

$$R - \overline{R} \cong F_R(\underline{X} - \overline{\underline{X}}). \qquad (5\ 31)$$

The range rate measurement becomes approximately modifiable as

$$\dot{R} - \overline{\dot{R}} \cong F_{RR}(\underline{X} - \overline{\underline{X}}) + F_{\dot{R}V}(\underline{V} - \overline{\underline{V}}). \qquad (5.32)$$

Define an error vector as (note the acceleration error is not included but can be done trivially)

$$e = (\underline{X} - \bar{\underline{X}}, \underline{V} - \bar{\underline{V}}, \underline{G} - \bar{\underline{G}}, \underline{K} - \bar{\underline{K}})^T \tag{5.33}$$

Then, the modifiable form is given by Eq. (5.7) where

$$F = F\,(z^*, \overline{\underline{X}}, \overline{\underline{V}}, \overline{\underline{G}}, \overline{\underline{K}})$$

$$= \begin{bmatrix} F_A & 0 & I & F_M \\ F_A & 0 & 0 & 0 \\ F_{RR} & F_{RV} & 0 & 0 \end{bmatrix} \tag{5.34}$$

and where the elements in the matrix are given in section 4.

5.6 Change in the Update Equation for Gain Calculation in the MGEKF

The update formulas for the EKF and the gain calculation are the same for the MGEKF, except that the update equation for the error variance equation when processing a measurement vector is changed from

$$P = (I - K_f h_x)\, M\, (I - K_f h_x) + K_f V K_f^T\,, \tag{5.35}$$

where K_f is the EKF gain, M is the a priori "error variance" before a measurement is processed, and P is the posteriori "error variance," to

$$P = (I - K_M F)\, M\, (I - K_M F) + K_M V K_M^T \tag{5.36}$$

where K_M is the gain of the MGEKF Therefore, the only change in the estimation algorithms is that h_x is replaced by F. Note that $F = f(z, \bar{\underline{X}}, \bar{\underline{V}}, \bar{\underline{G}}, \bar{\underline{K}})$, i.e., F is calculated using z rather than z*.

We now extend the above results to the case of distributed filtering for the SDI tracking problem. The structure of the MGEKF filter is similar to the KF for the linear case.

5.7 The Local MGEKF

Suppose that the suite of nonlinear measurements for sensor fusion are

$$z_j(k) = g_j(x(k)) + v_j(k) \qquad j = 1, \ldots, N(k) \tag{5.37}$$

where the noiseless measurement is $z_j^* = g_j(x)$ and $N(k)$ is the number of sensors that are to be combined at time stage k. The number may vary depending upon the purity of the current local estimators.

It is assumed that all the measurement functions are modifiable functions or at least approximately modifiable. Thus,

$$g_j(x) - g_j(\bar{x}) = H(z_j^*, \bar{x})\,(x - \bar{x}) \tag{5.38}$$

let us consider what this means for the local estimator equation

$$\hat{x}_j(k) = \bar{x}_j(k) + K_j(k)\,(z_j(k) - g_j(\bar{x}_j))\ , \tag{5.39}$$

$$\bar{x}_j(k) = \hat{x}_j(k \mid k-1)\ ,$$

where $K_j(k)$ is the gain of the MGEKF. Note that the estimator can be represented in terms of the modifiable function as

$$\hat{x}_j(k) = \bar{x}_j(k) + K_j(k)H(z_j^*, \bar{x}_j)\,(x(k) - \bar{x}_j(k)) + K_j(k)v_j(k) \tag{5.40}$$

Therefore, the error equations for $e_j(k) \stackrel{\Delta}{=} x(k) - \hat{x}_j(k)$ and $\bar{e}_j(k) \stackrel{\Delta}{=} x(k) - \bar{x}_j(k)$ are

$$e_j(k) = [I - K_j(k)H(z_j^*, \bar{x}_j)]\,\bar{e}_j(k) - K_j(k)v_j(k) \tag{5.41}$$

$$\bar{e}_j(k+1) = A(k)e_j(k) + w(k) \tag{5.42}$$

where (5.42) is obtained directly from the dynamic state equation. Note that without any approximation, assuming modifiable nonlinearities, the error equations (5.41) and (5.42)

are linear, even in the presence of noise. The essential improvement over the EKF is making use of this observation in constructing the gain algorithm for $K_j(k)$.

In constructing the gain algorithm one consideration was that the estimates should not be biased It is shown in Song and Speyer (1984) that if the gain is a function of the present measurement, then the estimator is highly biased. To avoid this, the gain was assumed using $\bar{H}_j(k) \,\Delta\, H_j(g_j(\bar{x}_j(k)), \bar{x}_j(k))$ to be of the form

$$K_j(k) = M_j(k)\bar{H}_j(k)\,[\bar{H}_j(k)^T M_j(k)\bar{H}_j(k) + R_j(k)\,]^{-1} \tag{5.43}$$

where we have used the notation $M_j(k) = P_j(k \mid k-1)$.In this way, $K_j(k)$ is only a function of past measurements. It should be noted that $\bar{H}_j(k)$ is essentially the partial of g_j with respect to x evaluated at $x = \bar{x}_j(k)$. In this way, the algorithm is still that of the EKF The difference arises in how the matrices P_j and M_j are propagated. These will be referred to as posterior and a prior error variances, but it must be remembered that these are not actually statistics and are really a kind of quasi-variance This statement also applies to the so-called statistical properties of the EKF. The error variances are propagated as

$$P_j(k) = [I - K_jH_j(z_j, \bar{x}_j)]M_j(k)[I - K_jH_j(z_j, \bar{x}_j)]^T + K_j(k)R_j(k)K_j(k)^T \tag{5.44}$$

$$M_j(k+1) = A(k)P_j(k)A(k)^T + Q(k) \tag{5.45}$$

where we have used the notation $P_j(k) = P_j(k \mid k)$ and $M_j(k+1) = P_j(k+1 \mid k)$. Note that $H_j(z_j, \bar{x}_j)$ is used in the update of $P_j(k)$ where the *actual* measurement is used. Furthermore, in the absence of measurement noise $P_j(k)$ reflects the actual error (5.41) whereas in the EKF the associated error variance does not reflect the actual error (5 41) Finally, it is shown in Song and Speyer (1984) that if z_j^* is used rather than z_j, then under reasonable assumptions the error is exponentially bounded in mean square.

5.8 The Global MGEKF

The local estimates are to be combined by the memoryless formula given in earlier chapters. Since the actual error is reflected by (5.44), it seems reasonable to combine the estimates according to the formula for the linear case where the effect of the error size is reflected by $P_j(k)$. At present we have no bounds on the quality of $P_j(k)$ in reflecting the true error variance, but determining these bounds is a subject of additional research. Since the EKF is based upon a local linearization, there is no reason to believe that the pseudo-error variance reflects the actual error variance unless the actual error is small. The question arises as to what estimate is best to perform the required local linearization. Since the MGEKF is based on a universal linearization and not a local linearization, this restriction does not apply.

In developing a global estimate, both P(k) and $h_j(k)$ need to be computed. The calculation of P(k) should be simply as

$$P(k)^{-1} = M(k)^{-1} + \sum_{j=1}^{N(k)} H_j(z_j, \bar{x}_j)^T R_j(k)^{-1} H(z_j, \bar{x}_j) \tag{5.46}$$

where M(k) is propagated using the same propagation formula as for the linear case. However, it is suggested that during periods where certain sensors are in doubt that this be calculated with respect to each of the suspect sensors eliminating the other suspect sensors For example, suppose that the IR sensor is affected by a flare and two signatures are being tracked It is suggested that (5.46) be calculated for the three situations each of the two tracks on the IR sensor and without the IR sensor If the IR sensor is eliminated, the global estimates are unaffected by the countermeasure. However, when one of the two IR tracks are identified as the flare, the information from the other IR track is now known to be correct signal, and this information is used to obtain a better estimate than when the IR

sensor is eliminated. The difficulty expressed here is due to the fact that the error variances cannot be computed off-line for the nonlinear estimation problem.

Once P(k) is obtained from (5.46) then $h_j(k)$ of (5.43) can be computed. Note that both F(k) and $G_j(k)$ require the calculation of the global covariances P(k) and M(k). It is again suggested that during periods when anomalies are present that $h_j(k)$ be calculated for each hypothesis and used only when the anomaly is known not to be present in the sensor data as determined by the detection and isolation scheme.

CHAPTER 6

SUMMARY AND FUTURE RECOMMENDATIONS

6.1 Summary and Findings

The main goal of this project is to solve the decentralized estimation problem in presence of the correlated sensor noise in a multisensor environment. This is a three phase project: Phase I, Phase II, and Phase III. In Phase I, the feasibility of the underlying techniques will be established. In Phase II, the the problem will be formulated in a more general setting and the techniques will be developed in details followed by a commercialization plan in Phase III. The work reported here is the outcome of the Phase I effort.

The subject problem is a very significant one in an SDI framework. Typically an SDI target is tracked by one or more sensors located on different orbiting satellites. The problem is to estimate the target state from all of these measurements which can be computed either using centralized or decentralized estimation techniques In a centralized technique, all the sensor measurements are transmitted to a central node and the state is estimated by utilizing all the information simultaneously. On the other hand, in a decentralized framework, the nodes process the observed data locally and then transmits the processed data to the central node where these are fused into the globally optimum estimate The decentralized estimation is superior to a centralized one from system survivability and computational considerations.

In a realistic situation, the sensor noises are correlated because these sensors are observing the same target through the same atmospheric medium. In this case, the decentralized estimation is more difficult than the case of the uncorrelated sensor noise. There has lately emerged a considerable amount of literature in the area of decentralized

estimation, but none of these has addressed the problem with the correlated sensor noise. In Phase I, we have addressed this issue and found a solution of a simpler version of the problem under the assumption that the local processors have the full knowledge of the global model and the system uncertainties The more general version of the problem with non-identical global and local models, will be analyzed in Phase II. A brief summary of this report follows.

An up to date available results in the area of decentralized estimation with uncorrelated sensor noise have been reviewed in Chapters 2 and 3 Chapter 2 is devoted to the continuous time case and Chapter 3 to the discrete time case These reviews have put the problem with the correlated sensor noise in the right perspective and exposed the difficulty in solving it. We have made some non-trivial extensions of the available results which have also been incorporated into these chapters

The general problem of decentralized estimation with non-identical global and local models has been presented in Chapter 2 We have followed the works of Willsky et al (1982) in developing this chapter. The main result is given in Equation (2.14) and is shown by block diagram of Figure 2.3. This shows that the globally optimal estimate can be constructed at a central node dynamically from the local estimates But the structure of the central node in this scheme is complicated. We have reduced this complexity in the scheme of Equation (2.15) and the corresponding block diagram of Figure 2.4. In this scheme each node must send 2 vectors - an optimal estimate of the state and another data dependent vector, to the central node. The structure of the central node as well as the transmission requirement has been reduced further in the scheme of Equation (2.16) and Figure 2.5. In this scheme some computational burden has been transferred from the central node to the local nodes and, as a result, each local node needs to send only one data dependent vector to the central node. However, in the process, the complexity of the local nodes increases. To remove this drawback, an improved scheme is presented in Equations

(2 17) and Figure 2.6. In this final scheme, there is no need of constructing an optimal filter explicitly at each node and the dynamical equation at that node can be initialized to zero This last scheme plays an important role in the development of the corresponding results for the case of correlated sensor noise in Chapter 4.

Chapter 3 is the mirror image of Chapter 2 but for the discrete time case. This chapter is included to expose some implementation details which are distinct from the continuous time case. The major results of this chapter are due to Speyer (1979) who solved the problem with uncorrelated sensor noise and identical global and local models. We have extended this result to include the case of dissimilar local and global models. The main result is given in Equations (3.15) and the corresponding block diagram of Figure 3.3 As in the continuous time case, each node must send 2 vectors to the central node. The transmission requirement and the complexity of the global and local nodes are reduced in the scheme of Equations (3.17) and Figure 3 4 In this scheme, there is no explicit Kalman filter at any node and each node needs to send only one data dependent vector to the central node.

The subject issue fo this project, i.e. the decentralized estimation problem with correlated sensor noise is presented in Chapter 4 under the assumption that the local processors have the full knowledge of the global model Because of the importance of this problem we have made this a self contained chapter. A brief review of the case of uncorrelated sensor noise is given in Section 4.2 The correlated sensor noise is dealt with in Section 4.3. This problem is first solved by transmitting all the sensor data to a central node and the results are given in Equation (4 10) Next, the technique of decorrelating the measurements from sensor to sensor is demonstrated in Equation (4 12) In this technique, the observed data is transformed into "new measurements" so that the data is uncorrelated across the sensors, but the solution depends upon the simultaneous availability of all the data The distributed algorithm is presented in Equation (4.13) and Figure 4.1. This is a

remarkably simple algorithm There is no explicit need of constructing a Kalman filter at the nodes and the initial conditions at the local nodes can be set to zero. Each local node must send to the global node only one data dependent vector the property of which is yet to be found out.

In many situations, particularly if the variables are expressed in the cartesian coordinates, the dynamics equations for the target state is linear but the measurement equations are nonlinear This class of problem is usually solved by using extended Kalman filters (EKF). In most of the cases, the implementation is based upon ad hoc expansions and linearization techniques. In this project, we have dealt with a special class of nonlinear measurements encountered in an SDI tracking problem known as modifiable nonlinearities The resulting filter is called the modifiable gain EKF (MGEKF) and has been dealt with in Chapter 5. In the case of bearings only measurements, it is anticipated that significant improvement in filter response and stability will result in the MGEKF formulation in contrast to that of the EKF. Since this is a new idea, this chapter has been written in a tutorial form The main result is given in Section 5.6 - 5.7. The decentralized version of the results are presented in Equations (5.38) - (5.45). These are the preliminary results and will be analyzed in detail in Phase II.

6.2 Future Recommendations

It is obvious from the earlier parts of this report that the decentralized estimation problem in presence of correlated sensor noise is indeed a very difficult one. We have solved a very simple version of this problem which assumes linear dynamics and measurement equations, identical global and local models, no computational and transmission delay, all nodes in broadcast mode etc. Therefore much work remains to be done which will form the basis of Phase II. The future recommendations include but are not necessarily limited to the following.

1. In a realistic environment, the local processors may not possibly have the complete knowledge of the global model Therefore the decentralized estimation problem in presence of the correlated sensor noise must be formulated for the case of dissimilar global and local models.

2. In Chapter 4, we have solved this problem only for the continuous time case. This analysis needs to be extended to the discrete time case

3. The IR sensors used in the SDI system typically provide information on bearings of the targets. Radar and laser trackers can provide additional information on range and range-rate In any event, if the motion of the target is described in a rectangular coordinate system, the above measurements are related nonlinearly to the state variables of the target. We have proposed in Chapter 5 that bearings only measurements can best be handled in a newly introduced framework called "modifiable functions." This is a very powerful technique and holds a great potential in solving the decentralized estimation problem with nonlinear measurements. So far the results are available for centralized estimation only and we have indicated in Chapter 5 how to proceed for the decentralized estimation. This problem needs to be formulated and analyzed in detail.

4 In a realistic situation, there is a delay in transmission from one node to the other and there is also a delay associated with the computation. In Phase I, we have assumed these delays to be zero. Time delays due to computation and communication processes must be incorporated in the decentralized estimation problem.

5. All the results of this report have been derived for one target only. This scenario needs to be broadened to multiple targets in future work

6 We have assumed that all the sensors are at fixed positions in an inertial frame while the target is moving with respect to this frame. But in an actual SDI scenario, the sensors are located in orbiting satellites and therefore have their own dynamics. Sensor dynamics must be included in the future formulation of the decentralized estimation problem.

7 We have assumed that the satellites carrying the sensors are communicating in a broadcast mode. But in practice, it is more likely that direct communications will take place only between adjacent satellites arranged in a ring structure about the Earth and communications will be allowed in either direction. The decentralized estimation problem needs to be formulated by incorporating the communication protocol of a ring structure

8. We have argued while establishing the superiority of the decentralized estimation over the centralized one that loss of the centralized processor could result in total or at least serious degradation of tracking capability. It will most likely turn out that distributed estimation will result in a more survivable and less complex tracking system It needs to be investigated how the system performance degrades in the event of one or more sensor failure. Two types of sensor failure must be considered. soft (degraded) failure and hard failure Soft failures can be modelled via use of statistical measurement noise models, while hard failures can be modelled by assuming complete loss of measurement.

9 Communication failure models should also be included in the decentralized estimation problem As in the case of sensor failure, communication failure can also be modelled as either hard or soft A hard failure, for example, could result from successful jamming or failure of the link to operate. Soft failures can be handled using statistical measurement noise models.

REFERENCES

Bryson, A.E. and Y.C. Ho, (1975), *Applied Optimal Control*, John Wiley and Sons, New York.

Castanon, D.A. and D. Teneketzis (1985), "Distributed Estimation Algorithms for Nonlinear Systems," *IEEE Trans on Automatic Control*, Vol. AC-30, No 5, May, pp 418-425.

Chang, T S. (1980), "A Stochastic Analysis of A Modified Gain Extended Kalman Filter with Applicatins to Estimation with Bearings Only Measurements," *IEEE Trans. on Auto Control*, Vol AC-30, No. 10, October.

Hashemipour, H.R., S Roy, and A J. Laub (1988), "Decentralized Structures for Parallel Kalman Filtering," *IEEE Trans on Auto Control*, Vol. AC-33, No. 1, January, pp 88-94

Jazwinski, A.H. (1970), *Stochastic Processes and Filtering Theory*, Academic Press, New York

Mehra, R.K. (1971), "A Comparison of Several Nonlinear Filters for Reentry Vehicle Tracking," IEEE Trans. on Auto. Control, Vol AC-16, No. 4, August, pp. 3077-319

Safonov, M.G. and M Athans, (1978) "Robustness and Computational Aspects of Nonlinear Stochastic Estimators and Regulators," *IEEE Trans on Automatic Control, Vol AC-23*, No. 4, August.

Song, T L and J.L. Speyer, (1984), "The Modified Gain Extended Kalman Filter and Parameter Identification in Linear Systems,' *Proceedings of the 1984 American Controls Conference*, June.

Song, T.L. and J L Speyer, (1985), "A Stochastic Analysis of a Modified Gain Extended Kalman Filter with Applications to Estimation with Bearings Only Measurements," *IEEE Trans on Automatic Control, Vol AC-30*, NO. 10, October.

Speyer, J.L (1979), "Computation and Transmission Requirements for a Decentralized Linear-Quadratic-Gaussian Control Problem,' *IEEE Trans. on Auto Control*, Vol AC-24, No 2, pp. 266-269.

Willsky, A.S., M.G. Belio, D.A. Castanon, B C. Levy, G.C. Verghese, (1982), 'Combining and Updating of Local Estimates and Regional Maps Along Sets of One-Dimensional Tracks," *IEEE Trans on Auto Control*, Vol. AC-277, No. 4, August, pp. 799-813.